Being Sensitive to Uncertainty!

Leon Arriola\textsuperscript{1} & James Hyman\textsuperscript{1}

\textsuperscript{1}Theoretical Division  
T5–Applied Mathematics and Plasma Physics  
Previously T7–Mathematical Modeling & Analysis  
Los Alamos National Laboratory

This work was carried out under the auspices of Los Alamos National Security, LLC (LANS), operator of the Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396 with the U.S. Department of Energy.

MTBI: Summer 2010
Forward problem (FP) takes nominal input parameters $p$ and produces the associated output solution $u$.

\begin{itemize}
  \item $A\vec{u} = \vec{b}$ (Linear System of Equations) $p \in \{a_{ij}, b_i\}$
  \item $A\vec{u} = \lambda\vec{u}$ (Eigenvalue Problem) $p \in \{a_{ij}\}$
  \item $\frac{d\vec{u}}{dt} = \vec{f}(\vec{u}, t; p)$, $\vec{u}(0) = \vec{u}_0$ (Initial Value Problem)
\end{itemize}
Forward problem (FP) takes nominal input parameters $p$ and produces the associated output solution $u$.

- $A\tilde{u} = \tilde{b}$ (Linear System of Equations) $p \in \{a_{ij}, b_i\}$
- $A\tilde{u} = \lambda\tilde{u}$ (Eigenvalue Problem) $p \in \{a_{ij}\}$
- $\frac{d\tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p), \tilde{u}(0) = \tilde{u}_0$ (Initial Value Problem)
Forward problem (FP) takes nominal input parameters $p$ and produces the associated output solution $u$.

- $A\tilde{u} = \vec{b}$  \hspace{1em} (Linear System of Equations) $p \in \{a_{ij}, b_i\}$
- $A\tilde{u} = \lambda\tilde{u}$  \hspace{1em} (Eigenvalue Problem) $p \in \{a_{ij}\}$
- $\frac{d\tilde{u}}{dt} = \vec{f}(\tilde{u}, t; p), \tilde{u}(0) = \tilde{u}_0$  \hspace{1em} (Initial Value Problem)
Forward problem (FP) takes nominal input parameters $p$ and produces the associated output solution $u$.

- $A\tilde{u} = \tilde{b}$ \quad (Linear System of Equations) $p \in \{a_{ij}, b_i\}$
- $A\tilde{u} = \lambda\tilde{u}$ \quad (Eigenvalue Problem) $p \in \{a_{ij}\}$
- $\frac{d\tilde{u}}{dt} = f(\tilde{u}, t; p), \quad \tilde{u}(0) = \tilde{u}_0$ \quad (Initial Value Problem)
Forward Sensitivity Analysis (FSA)

Forward sensitivity analysis (FSA) introduces perturbations to the input parameters, via $\delta p$ and quantifies the subsequent perturbations to the output solution via $\delta u$.

\[ A\tilde{u} = \tilde{b} \mapsto \left( A + \delta A \right) \left( \tilde{u} + \delta \tilde{u} \right) = \tilde{b} + \delta \tilde{b} \]

\[ A\tilde{u} = \lambda \tilde{u} \mapsto \left( A + \delta A \right) \left( \tilde{u} + \delta \tilde{u} \right) = (\lambda + \delta \lambda) \left( \tilde{u} + \delta \tilde{u} \right) \]

\[ \frac{d\tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p) \mapsto \frac{d[\tilde{u} + \delta \tilde{u}]}{dt} = \tilde{f}(\tilde{u} + \delta \tilde{u}, t; p + \delta p) \]
Forward Sensitivity Analysis (FSA)

Forward sensitivity analysis (FSA) introduces perturbations to the input parameters, via \( \delta p \) and quantifies the subsequent perturbations to the output solution via \( \delta u \).

\[
\begin{align*}
\text{Perturbation of Parameter} & \quad p + \delta p \\
\text{Forward Sensitivity Analysis} & \\
\text{Perturbation of Output} & \quad u + \delta u \quad \text{or} \\
& \quad \text{Function(al) } J(u + \delta u)
\end{align*}
\]

- \( \sim \widehat{u} = \widehat{b} \mapsto \left( \sim A + \delta \sim A \right) (\widehat{u} + \delta \widehat{u}) = \widehat{b} + \delta \widehat{b} \)
- \( \sim \widehat{u} = \lambda \widehat{u} \mapsto \left( \sim A + \delta \sim A \right) (\widehat{u} + \delta \widehat{u}) = (\lambda + \delta \lambda) (\widehat{u} + \delta \widehat{u}) \)
- \( \frac{d \widehat{u}}{dt} = \tilde{f}(\widehat{u}, t; p) \mapsto \frac{d[\widehat{u} + \delta \widehat{u}]}{dt} = \tilde{f}(\widehat{u} + \delta \widehat{u}, t; p + \delta p) \)
Forward Sensitivity Analysis (FSA)

Forward sensitivity analysis (FSA) introduces perturbations to the input parameters, via $\delta p$ and quantifies the subsequent perturbations to the output solution via $\delta u$.

- $A\tilde{u} = \tilde{b} \mapsto \left( A + \delta A \right) (\tilde{u} + \delta \tilde{u}) = \tilde{b} + \delta \tilde{b}$
- $A\tilde{u} = \lambda \tilde{u} \mapsto \left( A + \delta A \right) (\tilde{u} + \delta \tilde{u}) = (\lambda + \delta \lambda) (\tilde{u} + \delta \tilde{u})$
- $\frac{d\tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p) \mapsto \frac{d[\tilde{u} + \delta \tilde{u}]}{dt} = \tilde{f}(\tilde{u} + \delta \tilde{u}, t; p + \delta p)$
Forward Sensitivity Analysis (FSA)

Forward sensitivity analysis (FSA) introduces perturbations to the input parameters, via $\delta p$ and quantifies the subsequent perturbations to the output solution via $\delta u$.

- $A\vec{u} = \vec{b} \rightarrow (A + \delta A) (\vec{u} + \delta \vec{u}) = \vec{b} + \delta \vec{b}$

- $A\vec{u} = \lambda \vec{u} \rightarrow (A + \delta A) (\vec{u} + \delta \vec{u}) = (\lambda + \delta \lambda) (\vec{u} + \delta \vec{u})$

- $\frac{d\vec{u}}{dt} = \vec{f}(\vec{u}, t; p) \rightarrow \frac{d[\vec{u} + \delta \vec{u}]}{dt} = \vec{f}(\vec{u} + \delta \vec{u}, t; p + \delta p)$
Forward Sensitivity Analysis (FSA)

If the solution $u$ is differentiable in the parameters $p$:

- $\sim \ A \hat{u} = \vec{b} \mapsto A \frac{\partial \hat{u}}{\partial p} = \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \hat{u}$

- $\sim \ A \hat{u} = \lambda \hat{u} \mapsto A \frac{\partial \hat{u}}{\partial p} + \frac{\partial A}{\partial p} \hat{u} = \lambda \frac{\partial \hat{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \hat{u}$

- $\frac{d \hat{u}}{dt} = f(\hat{u}, t; p) \mapsto \frac{d}{dt} \left[ \frac{\partial \hat{u}}{\partial p} \right] = D_{\hat{u}} [f] \frac{\partial \hat{u}}{\partial p} + \frac{\partial f}{\partial p}$
If the solution $u$ is differentiable in the parameters $p$:

- $A\tilde{u} = \tilde{b} \mapsto A\frac{\partial \tilde{u}}{\partial p} = \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p}\tilde{u}$

- $A\tilde{u} = \lambda\tilde{u} \mapsto A\frac{\partial \tilde{u}}{\partial p} + \frac{\partial A}{\partial p}\tilde{u} = \lambda\frac{\partial \tilde{u}}{\partial p} + \frac{\partial \lambda}{\partial p}\tilde{u}$

- $\frac{d\tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p) \mapsto \frac{d}{dt}\left[\frac{\partial \tilde{u}}{\partial p}\right] = D_{\tilde{u}}[\tilde{f}]\frac{\partial \tilde{u}}{\partial p} + \frac{\partial \tilde{f}}{\partial p}$
Forward Sensitivity Analysis (FSA)

If the solution $u$ is differentiable in the parameters $p$:

$$A \tilde{u} = \tilde{b} \mapsto A \underaccent{\hat{\partial}} u = \frac{\partial b}{\partial p} - \underaccent{\hat{\partial}} A \underaccent{\hat{\partial}} u$$

$$A \tilde{u} = \lambda \tilde{u} \mapsto A \underaccent{\hat{\partial}} u + \frac{\partial A}{\partial p} \tilde{u} = \lambda \frac{\partial u}{\partial p} + \frac{\partial \lambda}{\partial p} \tilde{u}$$

$$\frac{d \tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p) \mapsto \frac{d}{dt} \left[ \frac{\partial \tilde{u}}{\partial p} \right] = \underaccent{\hat{\partial}} \tilde{u} \tilde{f} \frac{\partial u}{\partial p} + \frac{\partial \tilde{f}}{\partial p}$$
Forward Sensitivity Analysis (FSA)

If the solution $u$ is differentiable in the parameters $p$:

- $A\tilde{u} = \tilde{b} \mapsto A \frac{\partial \tilde{u}}{\partial p} = \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u}$

- $A\tilde{u} = \lambda \tilde{u} \mapsto A \frac{\partial \tilde{u}}{\partial p} + \frac{\partial A}{\partial p} \tilde{u} = \lambda \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \tilde{u}$

- $\frac{d\tilde{u}}{dt} = \tilde{f}(\tilde{u}, t; p) \mapsto \frac{d}{dt} \left[ \frac{\partial \tilde{u}}{\partial p} \right] = D_{\tilde{u}}[\tilde{f}] \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \tilde{f}}{\partial p}$
Uncertainties in the input parameters enter the model and produce uncertainty in the output.

The output isn’t just a single value but rather a PDF as well.
Uncertainties in the input parameters enter the model and produce uncertainty in the output.

The output isn’t just a single value but rather a PDF as well.
Combined distribution for parameters $p_1$ and $p_2$. 

Distribution of both $p_1$ and $p_2$
Define the *normalized sensitivity indexes* (SI):

\[ S_p := \lim_{\delta p \to 0} \frac{\left( \frac{\delta u}{u} \right)}{\left( \frac{\delta p}{p} \right)} = \frac{p}{u} \frac{\partial u}{\partial p} \quad u \neq 0 \]

If \( J(u) \) is a functional of \( u \) SI:

\[ S_{Jp} := \lim_{\delta p \to 0} \frac{\left( \frac{\delta J(u)}{J(u)} \right)}{\left( \frac{\delta p}{p} \right)} = \frac{p}{J(u)} \frac{\partial J(u)}{\partial p} \quad J(u) \neq 0 \]
Define the *normalized sensitivity indexes* (SI):

\[ S_p := \lim_{\delta p \to 0} \left( \frac{\delta u}{\delta p} \right) \left( \frac{p}{u \partial_p} \right) = \frac{p}{u} \frac{\partial u}{\partial p} \quad u \neq 0 \]

If \( J(u) \) is a functional of \( u \) SI:

\[ S_{Jp} := \lim_{\delta p \to 0} \left( \frac{\delta J(u)}{J(u)} \right) \left( \frac{p}{\delta p} \right) \left( \frac{\partial J(u)}{J(u)} \right) = \frac{p}{J(u)} \frac{\partial J(u)}{\partial p} \quad J(u) \neq 0 \]
Define the normalized sensitivity indexes (SI):

\[ S_p := \lim_{\delta p \to 0} \frac{\frac{\delta u}{u}}{\frac{\delta p}{p}} = \frac{p}{u} \frac{\partial u}{\partial p} \quad u \neq 0 \]

If \( J(u) \) is a functional of \( u \) SI:

\[ S_{Jp} := \lim_{\delta p \to 0} \frac{\frac{\delta J(u)}{J(u)}}{\frac{\delta p}{p}} = \frac{p}{J(u)} \frac{\partial J(u)}{\partial p} \quad J(u) \neq 0 \]
Define the *normalized sensitivity indexes* (SI):

\[
S_p := \lim_{\delta p \to 0} \left( \frac{\delta u}{\delta p} \right) = \frac{p}{u} \frac{\partial u}{\partial p} \quad u \neq 0
\]

If \( J(u) \) is a functional of \( u \) SI:

\[
S_{Jp} := \lim_{\delta p \to 0} \left( \frac{\delta J(u)}{\delta p} \right) = \frac{p}{J(u)} \frac{\partial J(u)}{\partial p} \quad J(u) \neq 0
\]
Consider the linear system of equations

$$A\tilde{u} = \tilde{b}$$

- Input parameters: $p \in \{a_{ij}, b_i\}$
- Output: $\tilde{u}$

$$A\tilde{u} = \tilde{b} \rightarrow A\frac{\partial \tilde{u}}{\partial p} = \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u}$$

$$\text{DDT!}: \quad A^{-1}A\frac{\partial \tilde{u}}{\partial p} = A^{-1} \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)$$
Consider the linear system of equations

\[ A \tilde{u} = \vec{b} \]

Input parameters: \( p \in \{a_{ij}, b_i\} \)

Output: \( \tilde{u} \)

\[ A \tilde{u} = \vec{b} \iff A \begin{array}{c} \vec{\partial u} \\ \vec{\partial p} \end{array} = \begin{array}{c} \vec{\partial b} \\ \vec{\partial p} \end{array} - \begin{array}{c} \vec{\partial A} \\ \vec{\partial p} \end{array} \tilde{u} \]

DDT!:

\[ A^{-1} \begin{array}{c} \vec{\partial u} \\ \vec{\partial p} \end{array} = A^{-1} \begin{array}{c} \vec{\partial b} \\ \vec{\partial p} \end{array} - A^{-1} \begin{array}{c} \vec{\partial A} \\ \vec{\partial p} \end{array} \tilde{u} \]
Consider the linear system of equations

\[ \sim \mathbf{A} \sim \mathbf{u} = \mathbf{b} \]

Input parameters: \( p \in \{ a_{ij}, b_i \} \)

Output: \( \mathbf{u} \)

\[ \sim \mathbf{A} \sim \mathbf{u} = \sim \mathbf{b} \rightarrow \sim \mathbf{A} \frac{\partial \mathbf{u}}{\partial p} = \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial \mathbf{A}}{\partial p} \sim \mathbf{u} \]

DDT!:

\[ \sim \mathbf{A}^{-1} \sim \mathbf{A} \frac{\partial \mathbf{u}}{\partial p} = \sim \mathbf{A}^{-1} \left( \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial \mathbf{A}}{\partial p} \sim \mathbf{u} \right) \]
Consider the linear system of equations

\[ A\tilde{u} = \tilde{b} \]

Input parameters: \( p \in \{ a_{ij}, b_i \} \)

Output: \( \tilde{u} \)

\[ A\tilde{u} = \tilde{b} \mapsto A\frac{\partial \tilde{u}}{\partial p} = \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \]

\[ \text{DDT!:} \quad A^{-1}A\frac{\partial \tilde{u}}{\partial p} = A^{-1} \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right) \]
Consider the linear system of equations

\[ A \tilde{u} = \tilde{b} \]

- Input parameters: \( p \in \{a_{ij}, b_i\} \)
- Output: \( \tilde{u} \)

\[ A \tilde{u} = \tilde{b} \iff A \frac{\partial \tilde{u}}{\partial p} = \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \]

\[ DDT!: \quad A^{-1} \frac{\partial \tilde{u}}{\partial p} = A^{-1} \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right) \]
Linear System of Equations

\[ A \sim_{N \times N} \begin{bmatrix} \frac{\partial \vec{u}}{\partial p} \\ \frac{\partial \vec{u}}{\partial p} \end{bmatrix} = \begin{bmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \end{bmatrix} \]

\[ A \sim_{N \times 1} \begin{bmatrix} \frac{\partial \vec{u}}{\partial p} \end{bmatrix} = \begin{bmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \end{bmatrix} \]

\[ \text{DDT!}: \quad \text{Something} \cdot A \sim_{N \times 1} \begin{bmatrix} \frac{\partial \vec{u}}{\partial p} \end{bmatrix} = \text{Something} \cdot \left( \begin{bmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \end{bmatrix} \right) \]
Linear System of Equations

\[
\begin{align*}
\begin{bmatrix} N \times 1 \\ N \times N \end{bmatrix} A \frac{\partial \tilde{u}}{\partial p} &= \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \\
\begin{bmatrix} N \times 1 \\ N \times N \end{bmatrix} A \frac{\partial \tilde{u}}{\partial p} &= \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u}
\end{align*}
\]

\[
\text{DDT!}: \quad \text{Something} \cdot \begin{bmatrix} M \times N \text{ or } 1 \times N \\ N \times 1 \end{bmatrix} A \frac{\partial \tilde{u}}{\partial p} = \text{Something} \cdot \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)
\]
Linear System of Equations

\[
A \sim \frac{\partial \mathbf{u}}{\partial p} = \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial A}{\partial p} \mathbf{u}
\]

\[
A \sim \frac{\partial \mathbf{u}}{\partial p} = \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial A}{\partial p} \mathbf{u}
\]

\[
M \times N \text{ or } 1 \times N
\]

\[
\text{DDT!}: \quad \text{Something} \cdot A \sim \frac{\partial \mathbf{u}}{\partial p} = \text{Something} \cdot \left( \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial A}{\partial p} \mathbf{u} \right)
\]

Leon Arriola & James Hyman
Linear System of Equations

- DDT!:
  \[
  \begin{aligned}
  \begin{pmatrix} 1 \\ \vdots \\ N \end{pmatrix} \cdot \begin{pmatrix} \nabla \vec{u} \\
  \frac{\partial A}{\partial p} \end{pmatrix} = \begin{pmatrix} \vec{v}^T \end{pmatrix} \begin{pmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \end{pmatrix}
  \end{aligned}
  \]

- What’s \( \vec{v} \)?
  - Answer: I don’t know yet!

- Notice that \( \vec{v}^T \cdot \begin{pmatrix} A \end{pmatrix} \) is an \( 1 \times N \) vector.

- Let \( \vec{c}^T := \begin{pmatrix} \vec{v}^T \end{pmatrix} A \) in which case \( \begin{pmatrix} A^T \end{pmatrix} \vec{v} = \vec{c} \) (Adjoint Problem)

- What’s \( \vec{c} \)?
  - Answer: I don’t know yet—but will shortly!

Leon Arriola & James Hyman
Linear System of Equations

$$\begin{align*}
\text{DDT!}: \quad & \begin{bmatrix} 1 \times N \\ \bar{v}^T \end{bmatrix} \cdot \begin{bmatrix} \partial \bar{u} \over \partial p \\ N \times 1 \end{bmatrix} \sim \begin{bmatrix} \bar{A} \over \partial p \\ 1 \times 1 \end{bmatrix} = \bar{v}^T \left( \frac{\partial \bar{b}}{\partial p} - \frac{\partial \bar{A}}{\partial p} \bar{u} \right) \\
\text{What's } \bar{v} \text{???} \\
\text{Answer: I don't know yet!} \\
\text{Notice that } \bar{v}^T \cdot \begin{bmatrix} \bar{A} \over \partial p \\ N \times N \end{bmatrix} \text{ is an } 1 \times N \text{ vector.} \\
\text{Let } \bar{c}^T := \bar{v}^T \bar{A} \text{ in which case } \bar{A}^T \bar{v} = \bar{c} \text{ (Adjoint Problem)} \\
\text{What's } \bar{c} \text{???} \\
\text{Answer: I don't know yet--but will shortly!}
\end{align*}$$
Linear System of Equations

- **DDT!**: \( \overline{v^T} \cdot \overline{A} \frac{\partial \overline{u}}{\partial \overline{p}} = \overline{v}^T \left( \frac{\partial \overline{b}}{\partial \overline{p}} - \frac{\partial \overline{A}}{\partial \overline{p}} \overline{u} \right) \)

- What’s \( \overline{v} \)?
  - Answer: I don’t know yet!

- Notice that \( \overline{v}^T \cdot \overline{A} \) is an \( 1 \times N \) vector.

- Let \( \overline{c}^T := \overline{v}^T \overline{A} \) in which case \( \overline{A}^T \overline{v} = \overline{c} \) (Adjoint Problem)

- What’s \( \overline{c} \)?
  - Answer: I don’t know yet—but will shortly!
Linear System of Equations

\[ \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \cdot A \begin{bmatrix} \partial \vec{u} / \partial p \end{bmatrix} = \begin{bmatrix} \partial \vec{b} / \partial p - \partial \vec{u} / \partial p \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \]

What’s \( \vec{v} \)???

Answer: I don’t know yet!

Notice that \( \vec{v}^T \cdot A \) is an \( 1 \times N \) vector.

Let \( \vec{c}^T := \vec{v}^T A \) in which case \( \vec{c}^T = \vec{c} \) (Adjoint Problem)

What’s \( \vec{c} \)???

Answer: I don’t know yet—but will shortly!
Linear System of Equations

- DDT!: \[ \begin{pmatrix} 1 \times N \\ \vec{v}^T \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial \vec{u}}{\partial p} \end{pmatrix} = \begin{pmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial A \sim}{\partial p} \vec{u} \end{pmatrix} \]

- What’s \( \vec{v} \)???
  - Answer: I don’t know yet!

- Notice that \( \vec{v}^T \cdot \sim A \) is an \( 1 \times N \) vector.

- Let \( \vec{c}^T := \vec{v}^T A \) in which case \( \sim A^T \vec{v} = \vec{c} \) (Adjoint Problem)

- What’s \( \vec{c} \)???
  - Answer: I don’t know yet–but will shortly!
### Linear System of Equations

**DDT!:**

\[
\begin{pmatrix}
\frac{\partial}{\partial p} & \frac{\partial}{\partial p}
\end{pmatrix}
\begin{pmatrix}
\vec{u}
\end{pmatrix}
\sim
\begin{pmatrix}
\vec{v}^T
\end{pmatrix}
\begin{pmatrix}
\vec{b}
\end{pmatrix}
- \begin{pmatrix}
\vec{A}
\end{pmatrix}
\sim
\begin{pmatrix}
\vec{u}
\end{pmatrix}
\]

- What’s \( \vec{v} \)?
  - Answer: I don’t know yet!

- Notice that \( \vec{v}^T \cdot \begin{pmatrix} \vec{A} \end{pmatrix} \sim \) is an \( 1 \times N \) vector.

- Let \( \vec{c}^T := \vec{v}^T \begin{pmatrix} \vec{A} \end{pmatrix} \sim \) in which case \( \begin{pmatrix} \vec{A} \end{pmatrix}^T \vec{v} = \vec{c} \) (Adjoint Problem)

- What’s \( \vec{c} \)?
  - Answer: I don’t know yet—but will shortly!

Leon Arriola & James Hyman
Linear System of Equations

- DDT!: \( \begin{bmatrix} 1 \times N \\ \vec{v}^T \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial \vec{u}}{\partial p} \\ N \times 1 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \frac{\partial \vec{b}}{\partial p} - \frac{\partial \vec{A}}{\partial p} \vec{u} \end{bmatrix} \\ 1 \times 1 \end{bmatrix} \)

- What’s \( \vec{v} \)???
  - Answer: I don’t know yet!

- Notice that \( \begin{bmatrix} \vec{v}^T \\ N \times N \end{bmatrix} \cdot \begin{bmatrix} \vec{A} \end{bmatrix} \) is an \( 1 \times N \) vector.

- Let \( \begin{bmatrix} \vec{c}^T \end{bmatrix} := \vec{v}^T \begin{bmatrix} \vec{A} \end{bmatrix} \) in which case \( \begin{bmatrix} \vec{c} \end{bmatrix} = \vec{c} \) (Adjoint Problem)

- What’s \( \vec{c} \)???
  - Answer: I don’t know yet–but will shortly!
Linear System of Equations

So $\mathbf{v}^T A \frac{\partial \mathbf{u}}{\partial p} \sim \partial \mathbf{v}^T \left( \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial \mathbf{A}}{\partial p} \mathbf{u} \right)$ becomes

$$\mathbf{c}^T \frac{\partial \mathbf{u}}{\partial p} = \mathbf{v}^T \left( \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial \mathbf{A}}{\partial p} \mathbf{u} \right)$$

Notice that

$$\mathbf{c}^T \frac{\partial \mathbf{u}}{\partial p} = \begin{pmatrix} c_1 & c_2 & \cdots & c_N \end{pmatrix} \begin{pmatrix} \frac{\partial u_1}{\partial p} \\ \frac{\partial u_2}{\partial p} \\ \vdots \\ \frac{\partial u_N}{\partial p} \end{pmatrix}$$
Linear System of Equations

- So $\vec{v}^T A \frac{\partial \vec{u}}{\partial p} \sim \frac{\partial b}{\partial p} - \frac{\partial A}{\partial p} \vec{u}$ becomes

  $$\vec{c}^T \frac{\partial \vec{u}}{\partial p} = \vec{v}^T \left( \frac{\partial b}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)$$

- Notice that

  $$\vec{c}^T \frac{\partial \vec{u}}{\partial p} = \begin{pmatrix} c_1 & c_2 & \cdots & c_N \end{pmatrix} \begin{pmatrix} \frac{\partial u_1}{\partial p} \\ \frac{\partial u_2}{\partial p} \\ \vdots \\ \frac{\partial u_N}{\partial p} \end{pmatrix}$$
Linear System of Equations

- DDT!:

$$\vec{c}^T \frac{\partial \vec{u}}{\partial p} = (c_1, c_2, \cdots, c_N) \begin{pmatrix} \frac{\partial u_1}{\partial p} \\ \frac{\partial u_2}{\partial p} \\ \vdots \\ \frac{\partial u_N}{\partial p} \end{pmatrix}$$

- DDT!:

$$\vec{c}_k^T \frac{\partial \vec{u}}{\partial p} = \begin{pmatrix} 0 & \cdots & 0 & \overset{k}{1} & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial u_1}{\partial p} \\ \frac{\partial u_2}{\partial p} \\ \vdots \\ \frac{\partial u_N}{\partial p} \end{pmatrix}$$
Linear System of Equations

- **DDT!:**

\[
\vec{c}^T \frac{\partial \vec{u}}{\partial p} = (c_1 \ c_2 \ \cdots \ c_N) \begin{pmatrix}
\frac{\partial u_1}{\partial p} \\
\frac{\partial u_2}{\partial p} \\
\vdots \\
\frac{\partial u_N}{\partial p}
\end{pmatrix}
\]

- **DDT!:**

\[
\vec{c}_k^T \frac{\partial \vec{u}}{\partial p} = \begin{pmatrix}
0 & \cdots & 0 & 1 & 0 & \cdots & 0
\end{pmatrix}
\begin{pmatrix}
\frac{\partial u_1}{\partial p} \\
\frac{\partial u_2}{\partial p} \\
\vdots \\
\frac{\partial u_N}{\partial p}
\end{pmatrix}
\]

Leon Arriola & James Hyman
In order to solve \( A \frac{\partial \vec{u}}{\partial p} = \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \),

Premultiply both sides by \( \vec{v}_k^T \) and define \( \vec{c}_k \) where

\[
\vec{c}_k^T = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}
\]

The final answer is...
In order to solve $A \frac{\partial \vec{u}}{\partial p} \sim \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u}$

Premultiply both sides by $\vec{v}_k^T$ and define $A^T \vec{v}_k \sim \vec{c}_k$ where

$$\vec{c}_k^T = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \text{kth column} \end{pmatrix}$$

The final answer is
In order to solve \[ A \frac{\partial \mathbf{u}}{\partial p} \sim \frac{\partial \mathbf{b}}{\partial p} - \frac{\partial \mathbf{A}}{\partial p} \mathbf{u} \]

Premultiply both sides by \( \mathbf{v}_k^T \) and define \( \mathbf{A}^T \mathbf{v}_k = \overline{\mathbf{c}}_k \) where

\[
\overline{\mathbf{c}}_k^T = \begin{pmatrix}
0 & \cdots & 0 & 1 & 0 & \cdots & 0
\end{pmatrix}
\]

The final answer is
Linear System of Equations

- **Forward Problem:**
  \[ A\vec{u} = \vec{b} \]

- **Adjoint Problem:**
  \[ A^T\vec{v}_k = \vec{c}_k \]

- **Forward Sensitivity**
  \[
  \frac{\partial u_k}{\partial p} = \vec{v}_k^T \left( \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)
  \]
Linear System of Equations

- **Forward Problem:**
  \[ A\tilde{u} = \tilde{b} \]

- **Adjoint Problem:**
  \[ A^T\tilde{v}_k = \tilde{c}_k \]

- **Forward Sensitivity**
  \[ \frac{\partial u_k}{\partial p} = \tilde{v}_k^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right) \]
Linear System of Equations

- **Forward Problem:**
  \[ A\tilde{u} = \tilde{b} \]

- **Adjoint Problem:**
  \[ A^T\tilde{v}_k = \tilde{c}_k \]

- **Forward Sensitivity**
  \[
  \frac{\partial u_k}{\partial p} = \tilde{v}_k^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)
  \]
Deterministic SIR Model

Consider a disease which, after some period of time, confers immunity or possibly death.

Divide the population into one of three distinct states:

- Susceptible: S
- Infected/Infectious: I
- Removed/Recovered: R
Consider a disease which, after some period of time, confers immunity or possibly death.

Divide the population into one of three distinct states:
- Susceptible: $S$
- Infected/Infectious: $I$
- Removed/Recovered: $R$
Consider a disease which, after some period of time, confers immunity or possibly death.

Divide the population into one of three distinct states:

- Susceptible: $S$
- Infected/Infectious: $I$
- Removed/Recovered: $R$
Consider a disease which, after some period of time, confers immunity or possibly death.

Divide the population into one of three distinct states:

- Susceptible: $S$
- Infected/Infectious: $I$
- Removed/Recovered: $R$
Consider a disease which, after some period of time, confers
immunity or possibly death.

Divide the population into one of three distinct states:
- Susceptible: $S$
- Infected/Infectious: $I$
- Removed/Recovered: $R$
Progression of an individual through these states can be schematically described by the directed graph\(^1\)

Commonly used deterministic SIR model:

\[
\begin{align*}
\frac{dS}{dt} &= -rSI \\
\frac{dI}{dt} &= rSI - \mu I \\
\frac{dR}{dt} &= \mu I.
\end{align*}
\]

\(^1\)Stochastic models use MCMC/DAM
Progression of an individual through these states can be schematically described by the directed graph:

Commonly used deterministic SIR model:

\[
\begin{align*}
\frac{dS}{dt} &= -rSI \\
\frac{dI}{dt} &= rSI - \mu I \\
\frac{dR}{dt} &= \mu I.
\end{align*}
\]

1Stochastic models use MCMC/DAM
Deterministic SIR Model

- Progression of an individual through these states can be schematically described by the directed graph\(^1\)

\[\begin{align*}
S &\xrightarrow{} I \\
I &\xrightarrow{} R \\
S &\xrightarrow{} R
\end{align*}\]

- Commonly used deterministic SIR model:

\[
\begin{align*}
\frac{dS}{dt} &= -rSI \\
\frac{dI}{dt} &= rSI - \mu I \\
\frac{dR}{dt} &= \mu I.
\end{align*}
\]

\(^1\)Stochastic models use MCMC/DAM

Leon Arriola & James Hyman
Numerical Solution of SIR Model

- Numerical solution where $r = 0.25$, $\mu = 0.0025$, $S_0 = 0.9$ and $I_0 = 0.1$. 

![Graph showing the numerical solution of the SIR model with $S(t)$, $I(t)$, and $R(t)$ over time.]
FSE of SIR Model

- FSE wrt parameters \( r \) and \( \mu \)

\[
\frac{d}{dt} \left[ \frac{\partial S}{\partial r} \right] = -rI \frac{\partial S}{\partial r} - rS \frac{\partial I}{\partial r} - SI
\]

\[
\frac{d}{dt} \left[ \frac{\partial S}{\partial \mu} \right] = -rI \frac{\partial S}{\partial \mu} - rS \frac{\partial I}{\partial \mu}
\]

\[
\frac{d}{dt} \left[ \frac{\partial I}{\partial r} \right] = rI \frac{\partial S}{\partial r} + [rS - \mu] \frac{\partial I}{\partial r} + SI
\]

\[
\frac{d}{dt} \left[ \frac{\partial I}{\partial \mu} \right] = rI \frac{\partial S}{\partial \mu} + [rS - \mu] \frac{\partial I}{\partial \mu} - I
\]

\[
\frac{d}{dt} \left[ \frac{\partial R}{\partial r} \right] = \mu \frac{\partial I}{\partial r}
\]

\[
\frac{d}{dt} \left[ \frac{\partial R}{\partial \mu} \right] = \mu \frac{\partial I}{\partial \mu} + I
\]
FSE of SIR Model

- FSE wrt parameters $r$ and $\mu$

\[
\frac{d}{dt} \left[ \frac{\partial S}{\partial r} \right] = -rI \frac{\partial S}{\partial r} - rS \frac{\partial I}{\partial r} - SI
\]

\[
\frac{d}{dt} \left[ \frac{\partial S}{\partial \mu} \right] = -rI \frac{\partial S}{\partial \mu} - rS \frac{\partial I}{\partial \mu}
\]

\[
\frac{d}{dt} \left[ \frac{\partial I}{\partial r} \right] = rI \frac{\partial S}{\partial r} + [rS - \mu] \frac{\partial I}{\partial r} + SI
\]

\[
\frac{d}{dt} \left[ \frac{\partial I}{\partial \mu} \right] = rI \frac{\partial S}{\partial \mu} + [rS - \mu] \frac{\partial I}{\partial \mu} - I
\]

\[
\frac{d}{dt} \left[ \frac{\partial R}{\partial r} \right] = \mu \frac{\partial I}{\partial r}
\]

\[
\frac{d}{dt} \left[ \frac{\partial R}{\partial \mu} \right] = \mu \frac{\partial I}{\partial \mu} + I
\]
What are the ICs?

Suppose that we want $\frac{\partial I}{\partial r}$

ICs are

$$\left. \frac{\partial I}{\partial r} \right|_{t=0} = 1$$

All others are set to zero
FSE of SIR Model

What are the ICs?

Suppose that we want $\partial I / \partial r$

ICs are

$\left. \frac{\partial I}{\partial r} \right|_{t=0} = 1$

All others are set to zero
What are the ICs?
Suppose that we want $\frac{\partial I}{\partial r}$
ICs are 

$$\left. \frac{\partial I}{\partial r} \right|_{t=0} = 1$$

All others are set to zero
What are the ICs?

Suppose that we want $\partial I/\partial r$

ICs are

$$\left. \frac{\partial I}{\partial r} \right|_{t=0} = 1$$

All others are set to zero
Time dependent sensitivity index of $I$ wrt $r$ and $\mu$

For $t \leq 30$, $I$ is most sensitive to changes in the parameter $r$, and almost unaffected for $t > 30$. 
• Time dependent sensitivity index of $I$ wrt $r$ and $\mu$

- For $t \leq 30$, $I$ is most sensitive to changes in the parameter $r$, and almost unaffected for $t > 30$. 

FSE of SIR Model

- FSE wrt initial conditions

\[
\frac{d}{dt}\left[ \frac{\partial S}{\partial S_0} \right] = -rI \frac{\partial S}{\partial S_0} - rS \frac{\partial I}{\partial S_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial S}{\partial I_0} \right] = -rI \frac{\partial S}{\partial I_0} - rS \frac{\partial I}{\partial I_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial S}{\partial R_0} \right] = -rI \frac{\partial S}{\partial R_0} - rS \frac{\partial I}{\partial R_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial I}{\partial S_0} \right] = rI \frac{\partial S}{\partial S_0} + [rS - \mu] \frac{\partial I}{\partial S_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial I}{\partial I_0} \right] = rI \frac{\partial S}{\partial I_0} + [rS - \mu] \frac{\partial I}{\partial I_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial I}{\partial R_0} \right] = rI \frac{\partial S}{\partial R_0} + [rS - \mu] \frac{\partial I}{\partial R_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial R}{\partial S_0} \right] = \mu \frac{\partial I}{\partial S_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial R}{\partial I_0} \right] = \mu \frac{\partial I}{\partial I_0}
\]

\[
\frac{d}{dt}\left[ \frac{\partial R}{\partial R_0} \right] = \mu \frac{\partial I}{\partial R_0}
\]
In order to calculate the sensitivity indexes, we first had to calculate the solutions to the system of three ODEs.

To do a full FSA, we must solve a total of 18 equations.

In modeling the chemical kinetics of certain reactions, it would not be unreasonable to have 10 equations with 20 parameters.

To do a full FSA would require solving a total of 310 odes.

Huge increase in the number of equations is a significant computational burden.
In order to calculate the sensitivity indexes, we first had to calculate the solutions to the system of three ODEs.

To do a full FSA, we must solve a total of 18 equations.

In modeling the chemical kinetics of certain reactions, it would not be unreasonable to have 10 equations with 20 parameters.

To do a full FSA would require solving a total of 310 odes.

Huge increase in the number of equations is a significant computational burden.
In order to calculate the sensitivity indexes, we first had to calculate the solutions to the system of three ODEs.

To do a full FSA, we must solve a total of 18 equations.

In modeling the chemical kinetics of certain reactions, it would not be unreasonable to have 10 equations with 20 parameters.

To do a full FSA would require solving a total of 310 odes.

Huge increase in the number of equations is a significant computational burden.
In order to calculate the sensitivity indexes, we first had to calculate the solutions to the system of three ODEs.

To do a full FSA, we must solve a total of 18 equations.

In modeling the chemical kinetics of certain reactions, it would not be unreasonable to have 10 equations with 20 parameters.

To do a full FSA would require solving a total of 310 odes.

Huge increase in the number of equations is a significant computational burden.
Proliferation of FSE’s

- In order to calculate the sensitivity indexes, we first had to calculate the solutions to the system of three ODEs.
- To do a full FSA, we must solve a total of 18 equations.
- In modeling the chemical kinetics of certain reactions, it would not be unreasonable to have 10 equations with 20 parameters.
- To do a full FSA would require solving a total of 310 odes.
- Huge increase in the number of equations is a significant computational burden.
Forward Sensitivity Analysis (FSA)

FSA is used when the number of output/solution variables of interest greatly exceeds the number of inputs/parameters.
ASA is used when the number of parameters/inputs of interest greatly exceeds the number of outputs/solutions.
Forward problem:

$$\frac{d\vec{u}}{dt} = \vec{F}[\vec{u}(t; \vec{p})], \quad \vec{u}(0) = \vec{u}_0$$

$\vec{u}$ is an $n \times 1$ forward solution vector and $\vec{p}$ is an $(k + n) \times 1$ vector which represents any of the $k$ parameters or $n$ initial conditions associated with the problem.

FSE

$$\frac{d}{dt} \left[ \tilde{D}_\vec{p} [\vec{u}] \right] = \tilde{D}_\vec{u} [\vec{F}] \cdot \tilde{D}_\vec{p} [\vec{u}] + \tilde{D}_\vec{p} [\vec{F}]$$

where $\tilde{D}$ are Jacobians.
Adjoint Sensitivity of Functionals for ODEs/IVP

- **Forward problem:**

  \[
  \frac{d\tilde{u}}{dt} = \tilde{F}[\tilde{u}(t; \tilde{p})], \quad \tilde{u}(0) = \tilde{u}_0
  \]

  \(\tilde{u}\) is an \(n \times 1\) forward solution vector and \(\tilde{p}\) is an \((k + n) \times 1\) vector which represents any of the \(k\) parameters or \(n\) initial conditions associated with the problem.

- **FSE**

  \[
  \frac{d}{dt} \left[ \sim_p [\tilde{u}] \right] = \sim_u [\tilde{F}] \cdot \sim_p [\tilde{u}] + \sim_p [\tilde{F}]
  \]

  where \(\sim\) are Jacobians.
Adjoint Sensitivity of Functionals for ODEs/IVP

- **Forward problem:**

\[
\frac{d\vec{u}}{dt} = \vec{F}[\vec{u}(t; \vec{p})], \quad \vec{u}(0) = \vec{u}_0
\]

\(\vec{u}\) is an \(n \times 1\) forward solution vector and \(\vec{p}\) is an \((k + n) \times 1\) vector which represents any of the \(k\) parameters or \(n\) initial conditions associated with the problem.

- **FSE**

\[
\frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] = D_{\vec{u}}[\vec{F}] \cdot D_{\vec{p}}[\vec{u}] + D_{\vec{p}}[\vec{F}]
\]

where \(D\) are Jacobians.
Adjoint Sensitivity of Functionals for ODEs/IVP

- **Forward problem:**

\[
\frac{d\tilde{u}}{dt} = \tilde{F}[\tilde{u}(t; \tilde{p})], \quad \tilde{u}(0) = \tilde{u}_0
\]

\(\tilde{u}\) is an \(n \times 1\) forward solution vector and \(\tilde{p}\) is an \((k + n) \times 1\) vector which represents any of the \(k\) parameters or \(n\) initial conditions associated with the problem.

- **FSE**

\[
\frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] = D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] + D_{\tilde{p}}[\tilde{F}]
\]

where \(D\) are Jacobians.
Adjoint Sensitivity of Functionals for ODEs/IVP

- **FSE** \( \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] = D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] + D_{\tilde{p}}[\tilde{F}] \)

Determine the sensitivity of an associated functional \( J(\tilde{u}) \) of the solution \( \tilde{u} \) where \( g \) and \( h \) are given scalar functions:

\[
J[\tilde{u}] := \left. \int_{t=0}^{b} g(\tilde{u}, \tilde{p}) \, dt + h(\tilde{u}, \tilde{p}) \right|_{t=b}
\]

- FSE for the functional \( J(\tilde{u}) \)

\[
\tilde{\nabla}_{\tilde{p}}[J] = \left. \int_{t=0}^{b} \left( D_{\tilde{p}}^{T}[\tilde{u}] \cdot \tilde{\nabla}_{\tilde{u}}[g] + \tilde{\nabla}_{\tilde{p}}[g] \right) \, dt \right|_{t=b} + \left. \left( D_{\tilde{p}}^{T}[\tilde{u}] \cdot \tilde{\nabla}_{\tilde{u}}[h] + \tilde{\nabla}_{\tilde{p}}[h] \right) \right|_{t=b}
\]
Adjoint Sensitivity of Functionals for ODEs/IVP

- FSE \( \frac{d}{dt} \left[ D_{\bar{p}} [\bar{u}] \right] = D_{\bar{u}} [\bar{F}] \cdot D_{\bar{p}} [\bar{u}] + D_{\bar{p}} [\bar{F}] \)

- Determine the sensitivity of an associated functional \( J(\bar{u}) \) of the solution \( \bar{u} \) where \( g \) and \( h \) are given scalar functions:

\[
J [\bar{u}] := \int_{t=0}^{b} g(\bar{u}, \bar{p}) \, dt + h(\bar{u}, \bar{p}) \bigg|_{t=b}
\]

- FSE for the functional \( J(\bar{u}) \)

\[
\nabla_{\bar{p}} [J] = \int_{t=0}^{b} \left( D_{\bar{p}}^T [\bar{u}] \cdot \nabla_{u} [g] + \nabla_{\bar{p}} [g] \right) \, dt
\]

\[
+ \left( D_{\bar{p}}^T [\bar{u}] \cdot \nabla_{u} [h] + \nabla_{\bar{p}} [h] \right) \bigg|_{t=b}
\]

Leon Arriola & James Hyman
Adjoint Sensitivity of Functionals for ODEs/IVP

- FSE \( \frac{d}{dt} \left[ D_{\bar{p}}[\bar{u}] \right] = D_{\bar{u}}[\bar{F}] \cdot D_{\bar{p}}[\bar{u}] + D_{\bar{p}}[\bar{F}] \)

Determine the sensitivity of an associated functional \( J(\bar{u}) \) of the solution \( \bar{u} \) where \( g \) and \( h \) are given scalar functions:

\[
J[\bar{u}] := \int_{t=0}^{b} g(\bar{u}, \bar{p}) \, dt + h(\bar{u}, \bar{p}) \bigg|_{t=b}
\]

- FSE for the functional \( J(\bar{u}) \)

\[
\hat{\nabla}_{\bar{p}}[J] = \int_{t=0}^{b} \left( D_{\bar{p}}^T[\bar{u}] \cdot \hat{\nabla}_{\bar{u}}[g] + \hat{\nabla}_{\bar{p}}[g] \right) \, dt
\]

\[+ \left( D_{\bar{p}}^T[\bar{u}] \cdot \hat{\nabla}_{\bar{u}}[h] + \hat{\nabla}_{\bar{p}}[h] \right) \bigg|_{t=b} \]
Adjoint Sensitivity of Functionals for ODEs/IVP

- FSE \[ \frac{d}{dt} \left[ \frac{D}{\sim \bar{p}} [\bar{u}] \right] = \frac{D}{\sim \bar{u}} [\bar{F}] \cdot \frac{D}{\sim \bar{p}} [\bar{u}] + \frac{D}{\sim \bar{p}} [\bar{F}] \]

- Determine the sensitivity of an associated functional \( J(\bar{u}) \) of the solution \( \bar{u} \) where \( g \) and \( h \) are given scalar functions:

\[
J [\bar{u}] := \int_{t=0}^{b} g(\bar{u}, \bar{p}) \, dt + h(\bar{u}, \bar{p}) \bigg|_{t=b}
\]

- FSE for the functional \( J(\bar{u}) \)

\[
\nabla_{\bar{p}} [J] = \int_{t=0}^{b} \left( \frac{D}{\sim \bar{p}} ^{T} [\bar{u}] \cdot \nabla_{\bar{u}} [g] + \nabla_{\bar{p}} [g] \right) \, dt
\]

\[
+ \left. \left( \frac{D}{\sim \bar{p}} ^{T} [\bar{u}] \cdot \nabla_{\bar{u}} [h] + \nabla_{\bar{p}} [h] \right) \right|_{t=b}
\]
Adjoint Sensitivity of Functionals for ODEs/IVP

- FSE \( \frac{d}{dt} \left[ \overleftarrow{D}_{\bar{p}}[\bar{u}] \right] = \overleftarrow{D}_u[F] \cdot \overleftarrow{D}_{\bar{p}}[\bar{u}] + \overleftarrow{D}_{\bar{p}}[\bar{F}] \)

- Determine the sensitivity of an associated functional \( J(\bar{u}) \) of the solution \( \bar{u} \) where \( g \) and \( h \) are given scalar functions:

\[
J[\bar{u}] := \int_{t=0}^{b} g(\bar{u}, \bar{p}) \, dt + h(\bar{u}, \bar{p}) \bigg|_{t=b}
\]

- FSE for the functional \( J(\bar{u}) \)

\[
\overleftarrow{\nabla}_{\bar{p}}[J] = \int_{t=0}^{b} \left( \overleftarrow{D}_{\bar{p}}^T[\bar{u}] \cdot \overleftarrow{\nabla}_u[g] + \overleftarrow{\nabla}_{\bar{p}}[g] \right) \, dt
\]

\[
+ \left( \overleftarrow{D}_{\bar{p}}^T[\bar{u}] \cdot \overleftarrow{\nabla}_u[h] + \overleftarrow{\nabla}_{\bar{p}}[h] \right) \bigg|_{t=b}
\]
Adjoint Sensitivity of Functionals for ODEs/IVP

\[ \vec{\nabla}_p [J] = \int_{t=0}^{b} \left( D_{\vec{p}}^T [\vec{u}] \cdot \vec{\nabla}_u [g] + \vec{\nabla}_p [g] \right) dt + \left( D_{\vec{p}}^T [\vec{u}] \cdot \vec{\nabla}_u [h] + \vec{\nabla}_p [h] \right) \bigg|_{t=b} \]

- We wish to eliminate having to directly calculate \( D_{\vec{p}}^T [\vec{u}] \)
- FSE \( \frac{d}{dt} \left[ D_{\vec{p}} [\vec{u}] \right] - D_{\vec{u}} [\vec{F}] \cdot D_{\vec{p}} [\vec{u}] - D_{\vec{p}} [\vec{F}] = 0 \)
- Define the standard inner product \( \langle A, \vec{b} \rangle := \int_{t=0}^{b} \vec{b}^T (t) \cdot A(t) dt \)
Adjoint Sensitivity of Functionals for ODEs/IVP

\[ \nabla_{\bar{p}}[J] = \int_{t=0}^{b} \left( D_{\bar{p}}^T[\bar{u}] \cdot \nabla_{\bar{u}}[g] + \nabla_{\bar{p}}[g] \right) dt + \left( D_{\bar{p}}^T[\bar{u}] \cdot \nabla_{\bar{u}}[h] + \nabla_{\bar{p}}[h] \right) \bigg|_{t=b} \]

- We wish to eliminate having to directly calculate \( D_{\bar{p}}^T[\bar{u}] \)
- FSE \( \frac{d}{dt} \left[ D_{\bar{p}}[\bar{u}] \right] - D_{\bar{u}}[\bar{F}] \cdot D_{\bar{p}}[\bar{u}] - D_{\bar{p}}[\bar{F}] = 0 \)
- Define the standard inner product \( \langle A, \bar{b} \rangle := \int_{t=0}^{b} \bar{b}^T(t) \cdot A(t) \, dt \)
Adjoint Sensitivity of Functionals for ODEs/IVP

\[ \nabla_{\vec{p}}[J] = \int_{t=0}^{b} \left( D_{\vec{p}}^T[\vec{u}] \cdot \nabla_{\vec{u}}[g] + \nabla_{\vec{p}}[g] \right) dt + \left( D_{\vec{p}}^T[\vec{u}] \cdot \nabla_{\vec{u}}[h] + \nabla_{\vec{p}}[h] \right) \bigg|_{t=b} \]

- We wish to eliminate having to directly calculate \( D_{\vec{p}}^T[\vec{u}] \)
- FSE \( \frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] - D_{\vec{u}}[\vec{F}] \cdot D_{\vec{p}}[\vec{u}] - D_{\vec{p}}[\vec{F}] = 0 \)
- Define the standard inner product \( \langle A, \vec{b} \rangle := \int_{t=0}^{b} \vec{b}^T(t) \cdot A(t) \, dt \)
Adjoint Sensitivity of Functionals for ODEs/IVP

Let \( \vec{v} \) be an unspecified adjoint variable and take the inner product

\[
\left\langle \frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] - D_{\vec{u}}[\vec{F}] \cdot D_{\vec{p}}[\vec{u}] - D_{\vec{p}}[\vec{F}], \vec{v} \right\rangle = \langle 0, \vec{v} \rangle = 0
\]

\[
\int_{t=0}^{b} \vec{v}^T \left( \frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] - D_{\vec{u}}[\vec{F}] \cdot D_{\vec{p}}[\vec{u}] - D_{\vec{p}}[\vec{F}] \right) dt = 0
\]

Derivative shift of \( \vec{v}^T \circ \frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] \) using integration by parts

\[
\vec{v}^T D_{\vec{p}}[\vec{u}] \bigg|_{t=0}^{b} + \int_{t=0}^{b} \left( -\frac{d\vec{v}^T}{dt} - \vec{v}^T D_{\vec{u}}[\vec{F}] \right) \frac{d}{dt} \left[ D_{\vec{p}}[\vec{u}] \right] dt - \int_{t=0}^{b} \vec{v}^T D_{\vec{p}}[\vec{F}] dt = 0
\]
Let $\tilde{v}$ be an unspecified adjoint variable and take the inner product

$$\left\langle \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] - D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] - D_{\tilde{p}}[\tilde{F}], \tilde{v} \right\rangle = \langle \tilde{0}, \tilde{v} \rangle = 0$$

$$\int_{t=0}^{b} \tilde{v}^T \left( \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] - D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] - D_{\tilde{p}}[\tilde{F}] \right) \, dt = 0$$

Derivative shift of $\tilde{v}^T \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right]$ using integration by parts

$$\tilde{v}^T D_{\tilde{p}}[\tilde{u}] \bigg|_{t=0}^{b} + \int_{t=0}^{b} \left( - \frac{d\tilde{v}^T}{dt} - \tilde{v}^T D_{\tilde{u}}[\tilde{F}] \right) D_{\tilde{p}}[\tilde{u}] \, dt - \int_{t=0}^{b} \tilde{v}^T D_{\tilde{p}}[\tilde{F}] \, dt = 0$$
Let $\vec{v}$ be an unspecified adjoint variable and take the inner product

$$\left\langle \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] - D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] - D_{\tilde{p}}[\tilde{F}], \vec{v} \right\rangle = \langle \vec{0}, \vec{v} \rangle = 0$$

$$\int_{t=0}^{b} \vec{v}^T \left( \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right] - D_{\tilde{u}}[\tilde{F}] \cdot D_{\tilde{p}}[\tilde{u}] - D_{\tilde{p}}[\tilde{F}] \right) \, dt = 0$$

Derivative shift of $\vec{v}^T \cdot \frac{d}{dt} \left[ D_{\tilde{p}}[\tilde{u}] \right]$ using integration by parts

$$\left. \vec{v}^T D_{\tilde{p}}[\tilde{u}] \right|_{t=0}^{b} + \int_{t=0}^{b} \left( - \frac{d\vec{v}^T}{dt} - \vec{v}^T D_{\tilde{u}}[\tilde{F}] \right) D_{\tilde{p}}[\tilde{u}] \, dt - \int_{t=0}^{b} \vec{v}^T D_{\tilde{p}}[\tilde{F}] \, dt = 0$$
Adjoint Sensitivity of Functionals for ODEs/IVP

- FS of the functional $J$ & inner product condition

$$\vec{\nabla}_p [J] = \left. \int_{t=0}^{b} \left( \vec{D}_p^T [\vec{u}] \cdot \vec{\nabla}_\vec{u} [g] + \vec{\nabla}_p [g] \right) \right|_{t=0}^b dt +$$

Compare this expression

$$\left. \left( \vec{D}_p^T [\vec{u}] \cdot \vec{\nabla}_\vec{u} [h] + \vec{\nabla}_p [h] \right) \right|_{t=b}$$

with this expression

$$\left. \vec{v}^T \vec{D}_p^T [\vec{u}] \right|_{t=0}^b + \int_{t=0}^b \left( -\frac{d\vec{v}^T}{dt} - \vec{v}^T \vec{D}_\vec{u}^T [\vec{F}] \right) \vec{D}_p^T [\vec{u}] dt - \int_{t=0}^b \vec{v}^T \vec{D}_p^T [\vec{F}] dt = 0$$
Adjoint Sensitivity of Functionals for ODEs/IVP

- Take transpose and compare terms

\[
\begin{align*}
\left( \left( -\frac{d\tilde{v}^T}{dt} - \tilde{v}^T D_{\vec{u}}[\vec{F}] \right) D_{\vec{p}}[\vec{u}] \right)^T &= D_{\vec{p}}^T[\vec{u}] \left( -\frac{d\tilde{v}}{dt} - D_{\vec{u}}^T[\vec{F}] \tilde{v} \right) \\
\upharpoonright & \\
D_{\vec{p}}^T[\vec{u}] \cdot \tilde{\nabla}_{\vec{u}}[g]
\end{align*}
\]

- Define the adjoint problem

\[
\frac{d\tilde{v}}{dt} + D_{\vec{u}}^T[\vec{F}] \tilde{v} := -\tilde{\nabla}_{\vec{u}}[g]
\]
Adjoint Sensitivity of Functionals for ODEs/IVP

- Take transpose and compare terms

\[
\left( \left( -\frac{d\vec{v}}{dt} - \vec{v}^T D_{\vec{u}} [\vec{F}] \right) D_{\vec{p}} [\vec{u}] \right)^T =
D_{\vec{p}}^T [\vec{u}] \left( -\frac{d\vec{v}}{dt} - D_{\vec{u}}^T [\vec{F}] \vec{v} \right)
\]

\[
\uparrow
D_{\vec{p}}^T [\vec{u}] \cdot \nabla_{\vec{u}} [g]
\]

- Define the adjoint problem

\[
\frac{d\vec{v}}{dt} + D_{\vec{u}}^T [\vec{F}] \vec{v} := -\nabla_{\vec{u}} [g]
\]
Take transpose and substitute

\[
\nabla_p[J] = \int_0^b \left( D_p^T[F]v + \nabla_p[g] \right) \, dt - \left. D_p^T[u]v \right|_{t=0}^b \\
+ \left. \left( D_p^T[u] \nabla_u[h] + \nabla_p[h] \right) \right|_{t=b}
\]

with adjoint problem

\[
\frac{d\tilde{v}}{dt} + D_{\tilde{u}}^T[F] \tilde{v} := -\nabla_{\tilde{u}}[g],
\]

and forward problem

\[
\frac{d\tilde{u}}{dt} = \tilde{F}[\tilde{u}(t; \tilde{p})], \quad \tilde{u}(0) = \tilde{u}_0
\]
Adjoint Sensitivity of Functionals for ODEs/IVP

- Take transpose and substitute

\[
\nabla_{\tilde{p}}[J] = \int_{t=0}^{b} \left( D_{\tilde{p}}^T [\tilde{F}] \tilde{v} + \nabla_{\tilde{p}}[g] \right) \, dt - D_{\tilde{p}}^T [\tilde{u}] \tilde{v} \bigg|_{t=0}^{b} \\
+ \left( D_{\tilde{p}}^T [\tilde{u}] \nabla_{\tilde{u}}[h] + \nabla_{\tilde{p}}[h] \right) \bigg|_{t=b}
\]

- with adjoint problem

\[
\frac{d\tilde{v}}{dt} + D_{\tilde{u}}^T [\tilde{F}] \tilde{v} := -\nabla_{\tilde{u}}[g],
\]

- and forward problem

\[
\frac{d\tilde{u}}{dt} = \tilde{F}[\tilde{u}(t; \tilde{p})], \quad \tilde{u}(0) = \tilde{u}_0
\]

Leon Arriola & James Hyman
Adjoint Sensitivity of Functionals for ODEs/IVP

- Take transpose and substitute

\[
\vec{\nabla}_p [J] = \int_{t=0}^{b} \left( D_{\sim p}^T [\vec{F}] \vec{v} + \vec{\nabla}_p [g] \right) \, dt - D_{\sim p}^T [\vec{u}] \vec{v} \bigg|_{t=0}^{b} \\
+ \left( D_{\sim p}^T [\vec{u}] \vec{\nabla}_u [h] + \vec{\nabla}_p [h] \right) \bigg|_{t=b}
\]

- with adjoint problem

\[
\frac{d\vec{v}}{dt} + D_{\sim \vec{u}}^T [\vec{F}] \vec{v} := -\vec{\nabla}_u [g],
\]

- and forward problem

\[
\frac{d\vec{u}}{dt} = \vec{F} [\vec{u}(t; \vec{p})], \quad \vec{u}(0) = \vec{u}_0
\]
In 1952, Harry Markowitz published a seminal paper titled “Portfolio Selection” which laid the foundation for what is now called modern portfolio theory.

Constructed the mathematical framework for the well known and accepted observation that investors, although seeking a maximum return on their investments, also simultaneously want to minimize the associated risk.

The proper mixture of various investments can significantly reduce the overall volatility of the portfolio, while maintaining a ”high” rate of return.

Quantitatively provide two solutions: a maximum amount of return for a given level of risk, or a minimum level of risk for a given amount of return.
In 1952, Harry Markowitz published a seminal paper titled “Portfolio Selection” which laid the foundation for what is now called modern portfolio theory.

Constructed the mathematical framework for the well known and accepted observation that investors, although seeking a maximum return on their investments, also simultaneously want to minimize the associated risk.

The proper mixture of various investments can significantly reduce the overall volatility of the portfolio, while maintaining a "high" rate of return.

Quantitatively provide two solutions: a maximum amount of return for a given level of risk, or a minimum level of risk for a given amount of return.
In 1952, Harry Markowitz published a seminal paper titled “Portfolio Selection” which laid the foundation for what is now called modern portfolio theory.

Constructed the mathematical framework for the well known and accepted observation that investors, although seeking a maximum return on their investments, also simultaneously want to minimize the associated risk.

The proper mixture of various investments can significantly reduce the overall volatility of the portfolio, while maintaining a ”high” rate of return.

Quantitatively provide two solutions: a maximum amount of return for a given level of risk, or a minimum level of risk for a given amount of return.
In 1952, Harry Markowitz published a seminal paper titled “Portfolio Selection” which laid the foundation for what is now called modern portfolio theory.

Constructed the mathematical framework for the well known and accepted observation that investors, although seeking a maximum return on their investments, also simultaneously want to minimize the associated risk.

The proper mixture of various investments can significantly reduce the overall volatility of the portfolio, while maintaining a "high" rate of return.

Quantitatively provide two solutions: a maximum amount of return for a given level of risk, or a minimum level of risk for a given amount of return.
Since cereal grains, such as wheat, provide a substantial portion of the caloric needs of humans worldwide, issues such as disease management and prevention are of the utmost importance.

Effects of soil type, average rainfall, disease tolerance, etc., on the yield, and hence the bottom line.

To further complicate the problem, agricultural researchers are attempting to produce perennial grain crops that will displace the annual crops that are currently planted.

The commonly used practices, that reduce disease inoculum in annual crops, such as tillage, delayed planting, or crop rotation, are not applicable to perennial crops.

Farmers would need to plant blends of seeds from a mixture of cultivars (varieties).
Since cereal grains, such as wheat, provide a substantial portion of the caloric needs of humans worldwide, issues such as disease management and prevention are of the utmost importance.

Effects of soil type, average rainfall, disease tolerance, etc., on the yield, and hence the bottom line.

To further complicate the problem, agricultural researchers are attempting to produce perennial grain crops that will displace the annual crops that are currently planted.

The commonly used practices, that reduce disease inoculum in annual crops, such as tillage, delayed planting, or crop rotation, are not applicable to perennial crops.

Farmers would need to plant blends of seeds from a mixture of cultivars (varieties).
Since cereal grains, such as wheat, provide a substantial portion of the caloric needs of humans worldwide, issues such as disease management and prevention are of the utmost importance.

Effects of soil type, average rainfall, disease tolerance, etc., on the yield, and hence the bottom line.

To further complicate the problem, agricultural researchers are attempting to produce perennial grain crops that will displace the annual crops that are currently planted.

The commonly used practices, that reduce disease inoculum in annual crops, such as tillage, delayed planting, or crop rotation, are not applicable to perennial crops.

Farmers would need to plant blends of seeds from a mixture of cultivars (varieties).
Wheat Selection

- Since cereal grains, such as wheat, provide a substantial portion of the caloric needs of humans worldwide, issues such as disease management and prevention are of the utmost importance.
- Effects of soil type, average rainfall, disease tolerance, etc., on the yield, and hence the bottom line.
- To further complicate the problem, agricultural researchers are attempting to produce perennial grain crops that will displace the annual crops that are currently planted.
- The commonly used practices, that reduce disease inoculum in annual crops, such as tillage, delayed planting, or crop rotation, are not applicable to perennial crops.
- Farmers would need to plant blends of seeds from a mixture of cultivars (varieties).
Since cereal grains, such as wheat, provide a substantial portion of the caloric needs of humans worldwide, issues such as disease management and prevention are of the utmost importance.

Effects of soil type, average rainfall, disease tolerance, etc., on the yield, and hence the bottom line.

To further complicate the problem, agricultural researchers are attempting to produce perennial grain crops that will displace the annual crops that are currently planted.

The commonly used practices, that reduce disease inoculum in annual crops, such as tillage, delayed planting, or crop rotation, are not applicable to perennial crops.

Farmers would need to plant blends of seeds from a mixture of cultivars (varieties).
Wheat Selection

- In the jargon of modern portfolio theory, investment in securities, stocks or bonds is replaced with the planting of multiple wheat cultivars.
- The objective of maximizing the expected rate of return on the investments is replaced with maximizing the wheat yield.
- Minimize the financial risks is replaced by minimizing the variation in wheat yield due to “genotype–environment interaction,” i.e., how each cultivar responds to the inevitable unpredictable environmental conditions.
- Risk is defined in terms of the standard deviation/variance of the return on the assets, and is in fact a quadratic functional.
- Once quantitative values can be established for the average yield, as well as the variance and covariance of yields of each cultivar, an optimal portfolio is found by solving a Quadratic Programming Problem (QPP).
In the jargon of modern portfolio theory, investment in securities, stocks or bonds is replaced with the planting of multiple wheat cultivars.

The objective of maximizing the expected rate of return on the investments is replaced with maximizing the wheat yield.

Minimize the financial risks is replaced by minimizing the variation in wheat yield due to “genotype–environment interaction,” i.e., how each cultivar responds to the inevitable unpredictable environmental conditions.

Risk is defined in terms of the standard deviation/variance of the return on the assets, and is in fact a quadratic functional.

Once quantitative values can be established for the average yield, as well as the variance and covariance of yields of each cultivar, an optimal portfolio is found by solving a Quadratic Programming Problem (QPP).
In the jargon of modern portfolio theory, investment in securities, stocks or bonds is replaced with the planting of multiple wheat cultivars.

The objective of maximizing the expected rate of return on the investments is replaced with maximizing the wheat yield.

Minimize the financial risks is replaced by minimizing the variation in wheat yield due to “genotype–environment interaction,” i.e., how each cultivar responds to the inevitable unpredictable environmental conditions.

Risk is defined in terms of the standard deviation/variance of the return on the assets, and is in fact a quadratic functional.

Once quantitative values can be established for the average yield, as well as the variance and covariance of yields of each cultivar, an optimal portfolio is found by solving a Quadratic Programming Problem (QPP).
Wheat Selection

- In the jargon of modern portfolio theory, investment in securities, stocks or bonds is replaced with the planting of multiple wheat cultivars.
- The objective of maximizing the expected rate of return on the investments is replaced with maximizing the wheat yield.
- Minimize the financial risks is replaced by minimizing the variation in wheat yield due to “genotype–environment interaction,” i.e., how each cultivar responds to the inevitable unpredictable environmental conditions.
- Risk is defined in terms of the standard deviation/variance of the return on the assets, and is in fact a quadratic functional.
- Once quantitative values can be established for the average yield, as well as the variance and covariance of yields of each cultivar, an optimal portfolio is found by solving a Quadratic Programming Problem (QPP).
In the jargon of modern portfolio theory, investment in securities, stocks or bonds is replaced with the planting of multiple wheat cultivars.

The objective of maximizing the expected rate of return on the investments is replaced with maximizing the wheat yield.

Minimize the financial risks is replaced by minimizing the variation in wheat yield due to “genotype–environment interaction,” i.e., how each cultivar responds to the inevitable unpredictable environmental conditions.

Risk is defined in terms of the standard deviation/variance of the return on the assets, and is in fact a quadratic functional.

Once quantitative values can be established for the average yield, as well as the variance and covariance of yields of each cultivar, an optimal portfolio is found by solving a **Quadratic Programming Problem (QPP)**.
Definition (QPP)

The QPP is defined as

\[
\text{Maximize } J(u_1, \ldots, u_n) := \vec{c}^T \vec{u} - \frac{1}{2} \vec{u}^T Q \vec{u},
\]

Subject to

\[
\begin{align*}
a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n & \leq b_1 \\
& \vdots \\
a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n & \leq b_m \\
u_1, \ldots, u_n & \geq 0
\end{align*}
\]

\(Q\) symmetric, positive semi--definite matrix

Leon Arriola & James Hyman

Los Alamos National Laboratory

EST, 1943
Quadratic Programming Problem

- Maximize the quadratic objective function

\[ J(u_1, \ldots, u_n) := \vec{c}^T \vec{u} - \frac{1}{2} \vec{u}^T \mathbf{Q} \vec{u} \]

- subject to the constraints

\[ \mathbf{A} \vec{u} \leq \vec{b} \]

- with nonnegativity conditions

\[ u_1, \ldots, u_n \geq 0 \]
Quadratic Programming Problem

- Maximize the quadratic objective function

\[ J(u_1, \ldots, u_n) := \vec{c}^T \vec{u} - \frac{1}{2} \vec{u}^T \mathbf{Q} \vec{u} \]

- subject to the constraints

\[ \mathbf{A} \vec{u} \leq \vec{b} \]

- with nonnegativity conditions

\[ u_1, \ldots, u_n \geq 0 \]
Quadratic Programming Problem

- Maximize the quadratic objective function
  \[ J(u_1, \ldots, u_n) := \vec{c}^T \vec{u} - \frac{1}{2} \vec{u}^T Q \vec{u} \]

- subject to the constraints
  \[ A\vec{u} \leq \vec{b} \]

- with nonnegativity conditions
  \[ u_1, \ldots, u_n \geq 0 \]
General Optimization Problem

- Maximize/Minimize a given objective function

\[ J(\vec{u}) = F(u_1, \ldots, u_n) \]

subject to the \( K \) equality and \( L \) inequality constraints

\[ f_k(\vec{u}) = 0 \quad \text{where} \quad k = 1, \ldots, K \]
\[ g_l(\vec{u}) \leq 0 \quad \text{where} \quad l = 1, \ldots, L. \]

- Define the modified Lagrangian function by forming a linear combination of the objective functional and the constraints as

\[ \mathcal{L}(\vec{u}; \mu, \lambda) := J(\vec{u}) + \sum_{k=1}^{K} \mu_k f_k(\vec{u}) + \sum_{l=1}^{L} \lambda_l g_l(\vec{u}), \]

where \( \mu_k \) and \( \lambda_l \) are called the Lagrange multipliers.

- The Lagrange multipliers are in fact adjoint variables.
General Optimization Problem

- Maximize/Minimize a given objective function

\[ J(\vec{u}) = F(u_1, \ldots, u_n) \]

subject to the \( K \) equality and \( L \) inequality constraints

\[ f_k(\vec{u}) = 0 \quad \text{where} \quad k = 1, \ldots, K \]
\[ g_l(\vec{u}) \leq 0 \quad \text{where} \quad l = 1, \ldots, L. \]

- Define the modified Lagrangian function by forming a linear combination of the objective functional and the constraints as

\[ \mathcal{L}(\vec{u}; \mu, \lambda) := J(\vec{u}) + \sum_{k=1}^{K} \mu_k f_k(\vec{u}) + \sum_{l=1}^{L} \lambda_l g_l(\vec{u}), \]

where \( \mu_k \) and \( \lambda_l \) are called the Lagrange multipliers.

- The Lagrange multipliers are in fact adjoint variables.
General Optimization Problem

- Maximize/Minimize a given objective function

\[ J(\vec{u}) = F(u_1, \ldots, u_n) \]

subject to the \( K \) equality and \( L \) inequality constraints

\[
\begin{align*}
  f_k(\vec{u}) &= 0 \quad \text{where} \quad k = 1, \ldots, K \\
  g_l(\vec{u}) &\leq 0 \quad \text{where} \quad l = 1, \ldots, L.
\end{align*}
\]

- Define the modified Lagrangian function by forming a linear combination of the objective functional and the constraints as

\[
\mathcal{L}(\vec{u}; \mu, \lambda) := J(\vec{u}) + \sum_{k=1}^{K} \mu_k f_k(\vec{u}) + \sum_{l=1}^{L} \lambda_l g_l(\vec{u}),
\]

where \( \mu_k \) and \( \lambda_l \) are called the Lagrange multipliers.
- The Lagrange multipliers are in fact adjoint variables.
Theorem (Karush/Kuhn/Tucker Theorem)

An optimal solution is found by solving the associated equations

\[
\frac{\partial J(\vec{u}^*)}{\partial u_j} + \sum_{k=1}^{K} \mu_k \frac{\partial f_k(\vec{u}^*)}{\partial u_j} + \sum_{l=1}^{L} \lambda_l \frac{\partial g_l(\vec{u}^*)}{\partial u_j} = 0 \quad \text{for } j = 1, \ldots n
\]

\[
\mu_k f_k(\vec{u}^*) = 0 \quad \text{for } k = 1, \ldots L
\]

\[
\lambda_l g_l(\vec{u}^*) = 0 \quad \text{for } l = 1, \ldots L
\]

where \( \vec{u}^* \) is the optimal solution.
Quadratic Programming Problem

- The inequality constraints are transformed into equality constraints by the introduction of slack variables.
- Construct the extended Lagrange function.

\[ \mathcal{L} := \bar{c}^T \bar{u} - \frac{1}{2} \bar{u}^T Q \bar{u} + \bar{v}^T \begin{pmatrix} b - A \bar{u} - \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} \end{pmatrix}. \]

- The optimal solution occurs at a critical point of the Lagrange function:

\[ \frac{\partial \mathcal{L}}{\partial u_j} = 0, \quad \frac{\partial \mathcal{L}}{\partial s_i} = 0, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial v_i} = 0. \]
Quadratic Programming Problem

- The inequality constraints are transformed into equality constraints by the introduction of slack variables.
- Construct the extended Lagrange function

\[
L := \bar{c}^T \bar{u} - \frac{1}{2} \bar{u}^T Q \bar{u} + \bar{v}^T \begin{pmatrix}
\bar{b} - \bar{A}\bar{u} - \\
(s_1)^2 \\
(s_2)^2 \\
\vdots \\
(s_m)^2
\end{pmatrix}.
\]

- The optimal solution occurs at a critical point of the Lagrange function:

\[
\frac{\partial L}{\partial u_j} = 0, \quad \frac{\partial L}{\partial s_i} = 0, \quad \text{and} \quad \frac{\partial L}{\partial v_i} = 0.
\]
Quadratic Programming Problem

- The inequality constraints are transformed into equality constraints by the introduction of slack variables.
- Construct the extended Lagrange function

\[ \mathcal{L} := \bar{c}^T \bar{u} - \frac{1}{2} \bar{u}^T Q \bar{u} + \bar{v}^T \begin{pmatrix} \bar{b} - A\bar{u} - \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} \end{pmatrix} . \]

- The optimal solution occurs at a critical point of the Lagrange function:

\[ \frac{\partial \mathcal{L}}{\partial u_j} = 0, \quad \frac{\partial \mathcal{L}}{\partial s_i} = 0, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial v_i} = 0. \]
Quadratic Programming Problem

- These equations respectively reduce to the mixed nonhomogeneous adjoint problem:

\[ A^T \vec{v} = \vec{c} - Q\vec{u}, \]

- the orthogonality conditions

\[ \nu_i s_i = 0, \quad \text{for} \quad i = 1, \ldots, m, \]

- and lastly to the forward problem

\[ A\vec{u} + \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} = \vec{b}. \]
These equations respectively reduce to the mixed nonhomogeneous adjoint problem:

\[ A^T \bar{v} = \bar{c} - Q \bar{u}, \]

the orthogonality conditions

\[ v_i s_i = 0, \quad \text{for} \quad i = 1, \ldots, m, \]

and lastly to the forward problem

\[ A \bar{u} + \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} = \bar{b}. \]
Quadratic Programming Problem

- These equations respectively reduce to the mixed nonhomogeneous adjoint problem:

\[ A^T \tilde{v} = \tilde{c} - Q\tilde{u}, \]

- the orthogonality conditions

\[ v_i s_i = 0, \quad \text{for} \quad i = 1, \ldots, m, \]

- and lastly to the forward problem

\[ A\tilde{u} + \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} = \tilde{b}. \]
Let $p$ denote any of the parameters $a_{ij}, b_i, c_j, \text{ or } q_{ij}$, where $q_{ij}$ denotes the $i,j$ entry of the matrix $Q$.

Differentiate the objective function, wrt parameter $p$:

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \frac{1}{2} \left( 2 \tilde{c}^T \frac{\partial \tilde{u}}{\partial p} - \tilde{u}^T \tilde{Q} \frac{\partial \tilde{u}}{\partial p} - \frac{\partial \tilde{u}^T}{\partial p} \tilde{Q} \tilde{u} \right)$$

Since the matrix $\tilde{Q}$ is symmetric, then

$$\left( \frac{\partial \tilde{u}}{\partial p} \right)^T = \frac{\partial \tilde{u}^T}{\partial p} \tilde{Q},$$

in which case $\frac{\partial J}{\partial p}$ reduces to

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \left( \tilde{c}^T - \tilde{u}^T \tilde{Q} \right) \frac{\partial \tilde{u}}{\partial p}.$$
Quadratic Programming Problem

- Let $p$ denote any of the parameters $a_{ij}, b_i, c_j,$ or $q_{ij}$, where $q_{ij}$ denotes the $i,j$ entry of the matrix $Q$.
- Differentiate the objective function, wrt parameter $p$:

\[
\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \frac{1}{2} \left( 2\tilde{c}^T \frac{\partial \tilde{u}}{\partial p} - \tilde{u}^T \tilde{Q} \frac{\partial \tilde{u}}{\partial p} - \frac{\partial \tilde{u}^T}{\partial p} \tilde{Q} \tilde{u} \right)
\]

- Since the matrix $\tilde{Q}$ is symmetric, then

\[
\left( \tilde{Q} \frac{\partial \tilde{u}}{\partial p} \right)^T = \frac{\partial \tilde{u}^T}{\partial p} \tilde{Q},
\]

in which case $\partial J/\partial p$ reduces to

\[
\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \left( \tilde{c}^T - \tilde{u}^T \tilde{Q} \right) \frac{\partial \tilde{u}}{\partial p}.
\]
Quadratic Programming Problem

Let $p$ denote any of the parameters $a_{ij}, b_i, c_j,$ or $q_{ij},$ where $q_{ij}$ denotes the $i,j$ entry of the matrix $Q$.

Differentiate the objective function, wrt parameter $p$:

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \frac{1}{2} \left( 2 \tilde{c}^T \frac{\partial \tilde{u}}{\partial p} - \tilde{u}^T Q \frac{\partial \tilde{u}}{\partial p} - \frac{\partial \tilde{u}^T}{\partial p} Q \tilde{u} \right)$$

Since the matrix $\tilde{Q}$ is symmetric, then

$$\left( \tilde{Q} \frac{\partial \tilde{u}}{\partial p} \right)^T = \frac{\partial \tilde{u}^T}{\partial p} Q,$$

in which case $\partial J/\partial p$ reduces to

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \left( \frac{\partial \tilde{Q}}{\partial p} \tilde{u}^T \right) \frac{\partial \tilde{u}}{\partial p}.$$
Quadratic Programming Problem

- Let $p$ denote any of the parameters $a_{ij}$, $b_i$, $c_j$, or $q_{ij}$, where $q_{ij}$ denotes the $i,j$ entry of the matrix $Q$.
- Differentiate the objective function, wrt parameter $p$:

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \bar{Q}}{\partial p} \tilde{u} + \frac{1}{2} \left( 2\tilde{c}^T \frac{\partial \tilde{u}}{\partial p} - \tilde{u}^T \bar{Q} \frac{\partial \tilde{u}}{\partial p} - \frac{\partial \tilde{u}^T}{\partial p} \bar{Q} \tilde{u} \right)$$

- Since the matrix $\bar{Q}$ is symmetric, then

$$\left( \bar{Q} \frac{\partial \tilde{u}}{\partial p} \right)^T = \frac{\partial \tilde{u}^T}{\partial p} \bar{Q},$$

- in which case $\frac{\partial J}{\partial p}$ reduces to

$$\frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \bar{Q}}{\partial p} \tilde{u} + \left( \tilde{c}^T - \tilde{u}^T \bar{Q} \right) \frac{\partial \tilde{u}}{\partial p}.$$
Quadratic Programming Problem

- The expression $\frac{\partial \vec{u}}{\partial p}$ will be replaced by an expression containing the forward and adjoint solutions.
  
  - This expression is found by differentiating the forward problem
    $\sim A \vec{u} + (s_1^2 s_2^2 \cdots s_m^2)^T = \vec{b}$ to get
    
    $$A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} + 2 \left(s_1 \frac{\partial s_1}{\partial p} s_2 \frac{\partial s_2}{\partial p} \cdots s_m \frac{\partial s_m}{\partial p}\right)^T = \frac{\partial \vec{b}}{\partial p}.$$

  - Next, premultiply this result by the adjoint solution $\vec{v}^T$ and use the orthogonality conditions $v_i s_i = 0$ to get
    
    $$\vec{v}^T A \frac{\partial \vec{u}}{\partial p} = \vec{v}^T \left(\frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u}\right)$$

  - in which case
    
    $$\frac{\partial J}{\partial p} = \frac{\partial c^T}{\partial p} \vec{u} - \frac{1}{2} \vec{u}^T \frac{\partial Q}{\partial p} \vec{u} + \vec{v}^T \left(\frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u}\right)$$
Quadratic Programming Problem

- The expression $\partial u / \partial p$ will be replaced by an expression containing the forward and adjoint solutions.
- This expression is found by differentiating the forward problem $A\tilde{u} + ((s_1)^2 (s_2)^2 \cdots (s_m)^2)^T = \tilde{b}$ to get
  
  \[ A \frac{\partial \tilde{u}}{\partial p} + \frac{\partial A}{\partial p} \tilde{u} + 2 \left( s_1 \frac{\partial s_1}{\partial p} \ s_2 \frac{\partial s_2}{\partial p} \cdots s_m \frac{\partial s_m}{\partial p} \right)^T = \frac{\partial \tilde{b}}{\partial p}. \]

- Next, premultiply this result by the adjoint solution $\tilde{v}^T$ and use the orthogonality conditions $v_i s_i = 0$ to get
  
  \[ \tilde{v}^T A \frac{\partial \tilde{u}}{\partial p} = \tilde{v}^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right) \]

- in which case
  
  \[ \frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial \tilde{Q}}{\partial p} \tilde{u} + \tilde{v}^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right) \]
Quadratic Programming Problem

- The expression $\partial \vec{u} / \partial p$ will be replaced by an expression containing the forward and adjoint solutions.
- This expression is found by differentiating the forward problem $A \vec{u} + \left( (s_1)^2 \ (s_2)^2 \ \cdots \ (s_m)^2 \right)^T = \vec{b}$ to get
  
  $$A \frac{\partial \vec{u}}{\partial p} \sim A \vec{u} + 2 \left( s_1 \frac{\partial s_1}{\partial p} \ s_2 \frac{\partial s_2}{\partial p} \ \cdots \ s_m \frac{\partial s_m}{\partial p} \right)^T = \frac{\partial \vec{b}}{\partial p}.$$ 

- Next, premultiply this result by the adjoint solution $\vec{v}^T$ and use the orthogonality conditions $v_i s_i = 0$ to get
  
  $$\vec{v}^T \frac{\partial \vec{u}}{\partial p} \sim \vec{v}^T \left( \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)$$

- in which case
  
  $$\frac{\partial J}{\partial p} = \frac{\partial \vec{c}^T}{\partial p} \vec{u} - \frac{1}{2} \vec{u}^T \frac{\partial Q}{\partial p} \vec{u} + \vec{v}^T \left( \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)$$
Quadratic Programming Problem

- The expression $\partial \vec{u} / \partial p$ will be replaced by an expression containing the forward and adjoint solutions.
- This expression is found by differentiating the forward problem $A \vec{u} + \left( (s_1)^2 \ (s_2)^2 \ \cdots \ (s_m)^2 \right)^T = \vec{b}$ to get
  
  $$A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} + 2 \left( s_1 \frac{\partial s_1}{\partial p} \ s_2 \frac{\partial s_2}{\partial p} \ \cdots \ s_m \frac{\partial s_m}{\partial p} \right)^T = \frac{\partial \vec{b}}{\partial p}.$$  

- Next, premultiply this result by the adjoint solution $\vec{v}^T$ and use the orthogonality conditions $v_i s_i = 0$ to get
  
  $$\vec{v}^T A \frac{\partial \vec{u}}{\partial p} = \vec{v}^T \left( \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)$$  

- in which case
  
  $$\frac{\partial J}{\partial p} = \frac{\partial \vec{c}^T}{\partial p} \vec{u} - \frac{1}{2} \vec{u}^T \frac{\partial Q}{\partial p} \vec{u} + \vec{v}^T \left( \frac{\partial \vec{b}}{\partial p} - \frac{\partial A}{\partial p} \vec{u} \right)$$
Quadratic Programming Problem—Summary

- **Forward problem:**
  \[
  A \tilde{u} + \begin{pmatrix}
  (s_1)^2 \\
  (s_2)^2 \\
  \vdots \\
  (s_m)^2
  \end{pmatrix} = \tilde{b}
  \]

- **Mixed nonhomogenous adjoint problem:**
  \[
  A^T \tilde{v} = \tilde{c} - Q\tilde{u}
  \]

- **Derivative of the objective functional:**
  \[
  \frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial Q}{\partial p} \tilde{u} + \tilde{v}^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)
  \]
Quadratic Programming Problem–Summary

- Forward problem:
  \[
  A\tilde{u} + \begin{pmatrix}
  (s_1)^2 \\
  (s_2)^2 \\
  . \\
  (s_m)^2
  \end{pmatrix} = \tilde{b}
  \]

- Mixed nonhomogenous adjoint problem:
  \[
  A^T\tilde{v} = \tilde{c} - Q\tilde{u}
  \]

- Derivative of the objective functional:
  \[
  \frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial Q}{\partial p} \tilde{u} + \tilde{v}^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)
  \]
Quadratic Programming Problem–Summary

- **Forward problem:**

  \[ A\tilde{u} + \begin{pmatrix} (s_1)^2 \\ (s_2)^2 \\ \vdots \\ (s_m)^2 \end{pmatrix} = \tilde{b} \]

- **Mixed nonhomogenous adjoint problem:**

  \[ A^T\tilde{v} = \tilde{c} - Q\tilde{u} \]

- **Derivative of the objective functional:**

  \[
  \frac{\partial J}{\partial p} = \frac{\partial \tilde{c}^T}{\partial p} \tilde{u} - \frac{1}{2} \tilde{u}^T \frac{\partial Q}{\partial p} \tilde{u} + \tilde{v}^T \left( \frac{\partial \tilde{b}}{\partial p} - \frac{\partial A}{\partial p} \tilde{u} \right)
  \]
Algorithmic Differentiation

**Annuity Function**

\[ f(m, l, r, t) = l \left( \frac{r}{m} \right) \left( 1 + \frac{r}{m} \right)^{mt} \frac{1}{\left( 1 + \frac{r}{m} \right)^{mt} - 1} \]

\( f \) returns the fixed periodic payment required to pay off a loan amount of \( l \), made for \( m \) periodic payments per year, with annual interest rate \( r \), and for a total of \( t \) years.

Loan of $100,000 is taken over 20 years, with annual interest of 15%, then the monthly payment is given by

\[ f(12, 100000, .15, 20) = 100000 \left( \frac{.15}{12} \right) \left( 1 + \frac{.15}{12} \right)^{12 \cdot 20} \frac{1}{\left( 1 + \frac{.15}{12} \right)^{12 \cdot 20} - 1} \]

\[ = 1316.80. \]
Algorithmic Differentiation

- Annuity Function

\[ f(m, l, r, t) = l \frac{\left( \frac{r}{m} \right) \left( 1 + \frac{r}{m} \right)^{mt}}{(1 + \frac{r}{m})^{mt} - 1} \]

- \( f \) returns the fixed periodic payment required to pay off a loan amount of \( l \), made for \( m \) periodic payments per year, with annual interest rate \( r \), and for a total of \( t \) years.

- Loan of $100,000 is taken over 20 years, with annual interest of 15%, then the monthly payment is given by

\[
\begin{align*}
\quad f(12, 100000, .15, 20) &= 100000 \frac{(.15 \frac{1}{12}) \left( 1 + .15 \frac{1}{12} \right)^{12 \cdot 20}}{(1 + .15 \frac{1}{12})^{12 \cdot 20} - 1} \\
&= 1316.80.
\end{align*}
\]
Algorithmic Differentiation

- **Annuity Function**

\[ f(m, l, r, t) = l \left( \frac{r}{m} \right) \left( 1 + \frac{r}{m} \right)^{mt} \left( 1 + \frac{r}{m} \right)^{mt} - 1 \]

- \( f \) returns the fixed periodic payment required to pay off a loan amount of \( l \), made for \( m \) periodic payments per year, with annual interest rate \( r \), and for a total of \( t \) years.

- Loan of $100,000 is taken over 20 years, with annual interest of 15%, then the monthly payment is given by

\[
 f(12, 100000, .15, 20) = 100000 \frac{(.15 \frac{12}{12}) \left( 1 + .15 \frac{12}{12} \right)^{12 \cdot 20}}{\left( 1 + .15 \frac{12}{12} \right)^{12 \cdot 20} - 1}
\]

\[ = 1316.80. \]
Algorithmic Differentiation

- **Annuity Function**

\[ f(m, l, r, t) = l \left( \frac{r}{m} \right) \left( 1 + \frac{r}{m} \right)^{mt} \left( 1 + \frac{r}{m} \right)^{-mt} - 1 \]

- \( f \) returns the fixed periodic payment required to pay off a loan amount of \( l \), made for \( m \) periodic payments per year, with annual interest rate \( r \), and for a total of \( t \) years.

- Loan of $100,000 is taken over 20 years, with annual interest of 15%, then the monthly payment is given by

\[
f(12, 100000, .15, 20) = 100000 \left( \frac{.15}{12} \right) \left( 1 + \frac{.15}{12} \right)^{12 \cdot 20} \left( 1 + \frac{.15}{12} \right)^{-12 \cdot 20} - 1 \]

\[= 1316.80.\]
Forward Evaluation Mode

- **Input variables are**

\[
\begin{align*}
p_1 & := m = 12.0 \\
p_2 & := l = 100000.0 \\
p_3 & := r = 0.15 \\
p_4 & := t = 20.0.
\end{align*}
\]

- **Intermediate variables**

\[
\begin{align*}
u_1 & := p_3 / p_1 = 0.0125 & r/m \\
u_2 & := 1 + u_1 = 1.0125 & 1 + r/m \\
u_3 & := p_1 p_4 = 240 & m \cdot t \\
u_4 & := u_1 u_2 u_3 = 0.2464 & (r/m)(1 + r/m)^{m \cdot t} \\
u_5 & := u_2 u_3 - 1 = 18.715 & (1 + r/m)^{m \cdot t} - 1 \\
u_6 & := u_4 / u_5 = 0.01368 \\
u_7 & := u_2 u_6 = 1316.79 & \text{Payment}
\end{align*}
\]

- **Output variable** \( u = u_7 = 1316.79 \)
Forward Evaluation Mode

- **Input variables are**

  \[ p_1 := m = 12.0 \]
  \[ p_2 := l = 100000.0 \]
  \[ p_3 := r = 0.15 \]
  \[ p_4 := t = 20.0. \]

- **Intermediate variables**

  \[ u_1 := p_3/p_1 = 0.0125 \]
  \[ u_2 := 1 + u_1 = 1.0125 \]
  \[ u_3 := p_1p_4 = 240 \]
  \[ u_4 := u_1u_2^{u_3} = 0.2464 \]
  \[ u_5 := u_2^{u_3} - 1 = 18.715 \]
  \[ u_6 := u_4/u_5 = 0.01368 \]
  \[ u_7 := u_2u_6 = 1316.79 \]

- **Output variable** \[ u = u_7 = 1316.79 \]
Forward Evaluation Mode

- **Input variables are**

  
  \[
  \begin{align*}
  p_1 & := m = 12.0 \\
  p_2 & := l = 100000.0 \\
  p_3 & := r = 0.15 \\
  p_4 & := t = 20.0.
  \end{align*}
  \]

- **Intermediate variables**

  \[
  \begin{align*}
  u_1 & := p_3/p_1 = 0.0125 \quad r/m \\
  u_2 & := 1 + u_1 = 1.0125 \quad 1 + r/m \\
  u_3 & := p_1p_4 = 240 \quad m \cdot t \\
  u_4 & := u_1u_2^{u_3} = 0.2464 \quad (r/m)(1 + r/m)^{m \cdot t} \\
  u_5 & := u_2^{u_3} - 1 = 18.715 \quad (1 + r/m)^{m \cdot t} - 1 \\
  u_6 & := u_4/u_5 = 0.01368 \\
  u_7 & := u_2u_6 = 1316.79 \quad \text{Payment}
  \end{align*}
  \]

- **Output variable** \( u = u_7 = 1316.79 \)
Forward Evaluation Mode

- Deterministic algorithm can be represented in a graphical format.

Abstract directed graph consists of two parts:
- Vertices represent the “objects”
- Directed edges represent “relationships”
Deterministic algorithm can be represented in a graphical format.

Abstract directed graph consists of two parts:
- Vertices represent the “objects”
- Directed edges represent ”relationships”
Deterministic algorithm can be represented in a graphical format.

Abstract directed graph consists of two parts:
- Vertices represent the “objects”
- Directed edges represent “relationships”
Forward Evaluation Mode

- Deterministic algorithm can be represented in a graphical format.

Abstract directed graph consists of two parts:
- Vertices represent the “objects”
- Directed edges represent ”relationships”
Forward Mode

- How does the payment change wrt changes in the interest rate?

\[
\frac{du}{dp_3} = \left( \frac{\partial u}{\partial u_7} \frac{\partial u_7}{\partial u_6} \frac{\partial u_6}{\partial u_4} \frac{\partial u_4}{\partial u_1} \frac{\partial u_1}{\partial p_3} \right) + \left( \frac{\partial u}{\partial u_7} \frac{\partial u_7}{\partial u_6} \frac{\partial u_6}{\partial u_4} \frac{\partial u_4}{\partial u_2} \frac{\partial u_2}{\partial u_1} \frac{\partial u_1}{\partial p_3} \right)
\]
Forward Mode

- How does the payment change wrt changes in the interest rate?

\[
\frac{du}{dp_3} = \frac{\partial u}{\partial u_7} \frac{\partial u_7}{\partial u_6} \frac{\partial u_6}{\partial u_4} \frac{\partial u_4}{\partial u_1} \frac{du_1}{dp_3} + \frac{\partial u}{\partial u_7} \frac{\partial u_7}{\partial u_6} \frac{\partial u_6}{\partial u_4} \frac{\partial u_4}{\partial u_2} \frac{\partial u_2}{\partial u_1} \frac{du_1}{dp_3}
\]
Due to precedence relations

\[ u_1 = u_1(p) \]
\[ u_2 = u_2(u_1, p) \]
\[ u_3 = u_3(u_2, u_1, p) \]
\[ \cdots \]
\[ u_N = u_N(u_{N-1}, u_{N-2}, \ldots, u_2, u_1, p) \]
Due to precedence relations

\[ u_1 = u_1(p) \]
General Forward Mode

Due to precedence relations

\[ u_1 = u_1(p) \]
\[ u_2 = u_2(u_1, p) \]
Due to precedence relations

\[ u_1 = u_1(p) \]
\[ u_2 = u_2(u_1, p) \]
\[ u_3 = u_3(u_2, u_1, p) \]
General Forward Mode

- Due to precedence relations

\[
\begin{align*}
    u_1 &= u_1(p) \\
    u_2 &= u_2(u_1, p) \\
    u_3 &= u_3(u_2, u_1, p) \\
    &\vdots
\end{align*}
\]
Due to precedence relations

\[ u_1 = u_1(p) \]
\[ u_2 = u_2(u_1, p) \]
\[ u_3 = u_3(u_2, u_1, p) \]
\[ \vdots \]
\[ u_N = u_N(u_{N-1}, u_{N-2}, \ldots, u_2, u_1, p) \]
General Forward Sensitivity Mode

\[
\frac{du_1}{dp} = \frac{\partial u_1}{\partial p}
\]
General Forward Sensitivity Mode

\[
\frac{du_1}{dp} = \frac{\partial u_1}{\partial p}
\]

\[
\frac{du_2}{dp} = \frac{\partial u_2}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_2}{\partial p}
\]
General Forward Sensitivity Mode

\[
\begin{align*}
\frac{du_1}{dp} &= \frac{\partial u_1}{\partial p} \\
\frac{du_2}{dp} &= \frac{\partial u_2}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_2}{\partial p} \\
\frac{du_3}{dp} &= \frac{\partial u_3}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_3}{\partial u_2} \frac{du_2}{dp} + \frac{\partial u_3}{\partial p}
\end{align*}
\]
General Forward Sensitivity Mode

\[
\frac{du_1}{dp} = \frac{\partial u_1}{\partial p}
\]
\[
\frac{du_2}{dp} = \frac{\partial u_2}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_2}{\partial p}
\]
\[
\frac{du_3}{dp} = \frac{\partial u_3}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_3}{\partial u_2} \frac{du_2}{dp} + \frac{\partial u_3}{\partial p}
\]

\vdots
General Forward Sensitivity Mode

\[
\frac{du_1}{dp} = \frac{\partial u_1}{\partial p}
\]

\[
\frac{du_2}{dp} = \frac{\partial u_2}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_2}{\partial p}
\]

\[
\frac{du_3}{dp} = \frac{\partial u_3}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_3}{\partial u_2} \frac{du_2}{dp} + \frac{\partial u_3}{\partial p}
\]

\[
\vdots
\]

\[
\frac{du_N}{dp} = \frac{\partial u_N}{\partial u_1} \frac{du_1}{dp} + \frac{\partial u_N}{\partial u_2} \frac{du_2}{dp} + \cdots + \frac{\partial u_N}{\partial u_{N-1}} \frac{du_{N-1}}{dp} + \frac{\partial u_N}{\partial p}
\]
This linear system can be written in the more concise form

\[
\left( D[\vec{u}] - 2I \right) \frac{d\vec{u}}{dp} = -\frac{\partial \vec{u}}{\partial p}
\]

where

\[
D[\vec{u}] = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
\frac{\partial u_2}{\partial u_1} & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\frac{\partial u_N}{\partial u_1} & \frac{\partial u_N}{\partial u_2} & \cdots & \frac{\partial u_N}{\partial u_{N-1}} & 1
\end{pmatrix}
\]

and

\[
\vec{u} = \begin{pmatrix}
u_1 \\
u_2 \\
\vdots \\
u_N
\end{pmatrix}
\]
Reverse Mode
Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_2} = \frac{\partial u}{\partial u_4} \frac{\partial u_4}{\partial u_2} + \frac{\partial u}{\partial u_5} \frac{\partial u_5}{\partial u_2} \\
\frac{\partial u}{\partial u_4} = \frac{\partial u}{\partial u_6} \frac{\partial u_6}{\partial u_4}
\]

\[
\frac{\partial u}{\partial u_2} = \frac{\partial u}{\partial u_1} \frac{\partial u_1}{\partial u_2} + \frac{\partial u}{\partial u_4} \frac{\partial u_4}{\partial u_1} \\
\frac{\partial u}{\partial u_5} = \frac{\partial u}{\partial u_6} \frac{\partial u_6}{\partial u_5}
\]

\[
\frac{\partial u}{\partial u_6} = \frac{\partial u}{\partial u_7} \frac{\partial u_7}{\partial u_6}
\]

\[
\frac{du}{dp_3} = \frac{\partial u}{\partial u_1} \frac{\partial u_1}{\partial p}
\]
General Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_N} = \frac{\partial u_N}{\partial u_N} = 1
\]
General Reverse Sensitivity Mode

\[ \frac{\partial u}{\partial u_N} = \frac{\partial u_N}{\partial u_N} = 1 \]

\[ \frac{\partial u}{\partial u_{N-1}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-1}} \]
Leon Arriola & James Hyman

General Reverse Sensitivity Mode

\[
\begin{align*}
\frac{\partial u}{\partial u_N} &= \frac{\partial u_N}{\partial u_N} = 1 \\
\frac{\partial u}{\partial u_{N-1}} &= \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-1}} \\
\frac{\partial u}{\partial u_{N-2}} &= \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-2}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-2}}
\end{align*}
\]
General Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_N} = \frac{\partial u_N}{\partial u_N} = 1
\]

\[
\frac{\partial u}{\partial u_{N-1}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-1}}
\]

\[
\frac{\partial u}{\partial u_{N-2}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-2}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-2}}
\]

\[
\frac{\partial u}{\partial u_{N-3}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-3}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-3}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-3}}
\]
General Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_N} = \frac{\partial u_N}{\partial u_N} = 1
\]

\[
\frac{\partial u}{\partial u_{N-1}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-1}}
\]

\[
\frac{\partial u}{\partial u_{N-2}} = \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-2}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-2}}
\]

\[
\frac{\partial u}{\partial u_{N-3}} = \frac{\partial u}{\partial u_{N-2}} \frac{\partial u_{N-2}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-3}}
\]

\[\vdots\]
General Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_N} = \frac{\partial u_N}{\partial u_N} = 1
\]

\[
\frac{\partial u}{\partial u_{N-1}} = \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-1}} \]

\[
\frac{\partial u}{\partial u_{N-2}} = \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-2}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-2}} \]

\[
\frac{\partial u}{\partial u_{N-3}} = \frac{\partial u}{\partial u_{N-2}} \frac{\partial u_{N-2}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_{N-3}} \]

\[
\vdots
\]

\[
\frac{\partial u}{\partial u_1} = \frac{\partial u}{\partial u_2} \frac{\partial u_2}{\partial u_1} + \frac{\partial u}{\partial u_3} \frac{\partial u_3}{\partial u_1} + \cdots + \frac{\partial u}{\partial u_N} \frac{\partial u_N}{\partial u_1}
\]
General Reverse Sensitivity Mode

\[
\frac{\partial u}{\partial u_{N}} = \frac{\partial u_{N}}{\partial u_{N}} = 1 \\
\frac{\partial u}{\partial u_{N-1}} = \frac{\partial u}{\partial u_{N}} \frac{\partial u_{N}}{\partial u_{N-1}} \\
\frac{\partial u}{\partial u_{N-2}} = \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-2}} + \frac{\partial u}{\partial u_{N}} \frac{\partial u_{N}}{\partial u_{N-2}} \\
\frac{\partial u}{\partial u_{N-3}} = \frac{\partial u}{\partial u_{N-2}} \frac{\partial u_{N-2}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_{N-1}} \frac{\partial u_{N-1}}{\partial u_{N-3}} + \frac{\partial u}{\partial u_{N}} \frac{\partial u_{N}}{\partial u_{N-3}} \\
\vdots \\
\frac{\partial u}{\partial u_{1}} = \frac{\partial u}{\partial u_{2}} \frac{\partial u_{2}}{\partial u_{1}} + \frac{\partial u}{\partial u_{3}} \frac{\partial u_{3}}{\partial u_{1}} + \cdots + \frac{\partial u}{\partial u_{N}} \frac{\partial u_{N}}{\partial u_{1}} \\
\frac{du}{dp} = \sum_{i=1}^{N} \frac{\partial u}{\partial u_{i}} \frac{\partial u_{i}}{dp}
\]
Eigenvalue Problem

- Consider the right eigenvalue problem
  \[ A\vec{u} = \lambda \vec{u} \]

- Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.

- Hence we have \( n \) linearly independent eigenvectors \( \vec{u}_k \).

- Input parameters \( p \in \{a_{ij}\} \)

- Outputs: \( \lambda_i, \vec{u}_i \)

- FSEs
  \[ A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} = \lambda \frac{\partial \vec{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \vec{u} \]

- This equation has two unknowns \( \partial\lambda/\partial a_{ij} \) and \( \partial\vec{u}/\partial a_{ij} \)

- Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Consider the right eigenvalue problem

\[ A \tilde{u} = \lambda \tilde{u} \]

Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.

Hence we have \( n \) linearly independent eigenvectors \( \tilde{u}_k \).

Input parameters \( p \in \{a_{ij}\} \)

Outputs: \( \lambda_i, \tilde{u}_i \)

FSEs

\[ A \frac{\partial \tilde{u}}{\partial p} + \frac{\partial A}{\partial p} \tilde{u} = \lambda \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \tilde{u} \]

This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \tilde{u} / \partial a_{ij} \)

Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Eigenvalue Problem

- Consider the right eigenvalue problem

\[ A \tilde{u} = \lambda \tilde{u} \]

- Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.

- Hence we have \( n \) linearly independent eigenvectors \( \tilde{u}_k \).

- Input parameters \( p \in \{a_{ij}\} \)

- Outputs: \( \lambda_i, \tilde{u}_i \)

- FSEs

\[ A \frac{\partial \tilde{u}}{\partial p} + \frac{\partial A}{\partial p} \tilde{u} = \lambda \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \tilde{u} \]

- This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \tilde{u} / \partial a_{ij} \)

- Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Eigenvalue Problem

- Consider the right eigenvalue problem

\[ \tilde{A}\tilde{u} = \lambda\tilde{u} \]

- Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.
- Hence we have \( n \) linearly independent eigenvectors \( \tilde{u}_k \).
- Input parameters \( p \in \{a_{ij}\} \)
- Outputs: \( \lambda_i, \tilde{u}_i \)
- FSEs

\[ \tilde{A} \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \tilde{A}}{\partial p} \tilde{u} = \lambda \frac{\partial \tilde{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \tilde{u} \]

- This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \tilde{u} / \partial a_{ij} \)
- Either obtain another independent equation or eliminate one of the unknown variables from this equation.

Leon Arriola & James Hyman
Consider the right eigenvalue problem

\[ A \vec{u} = \lambda \vec{u} \]

Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.
Hence we have \( n \) linearly independent eigenvectors \( \vec{u}_k \).

Input parameters \( p \in \{a_{ij}\} \)

Outputs: \( \lambda_i, \vec{u}_i \)

FSEs

\[ A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} = \lambda \frac{\partial \vec{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \vec{u} \]

This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \vec{u} / \partial a_{ij} \).

Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Consider the right eigenvalue problem

\[ A\vec{u} = \lambda \vec{u} \]

Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.

Hence we have \( n \) linearly independent eigenvectors \( \vec{u}_k \).

Input parameters \( p \in \{a_{ij}\} \)

Outputs: \( \lambda_i, \vec{u}_i \)

FSEs

\[ A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} = \lambda \frac{\partial \vec{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \vec{u} \]

This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \vec{u} / \partial a_{ij} \)

Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Consider the right eigenvalue problem

\[ A \vec{u} = \lambda \vec{u} \]

Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.

Hence we have \( n \) linearly independent eigenvectors \( \vec{u}_k \).

Input parameters \( p \in \{a_{ij}\} \)

Outputs: \( \lambda_i, \vec{u}_i \)

FSEs

\[ A \frac{\partial \vec{u}}{\partial p} + \frac{\partial A}{\partial p} \vec{u} = \lambda \frac{\partial \vec{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \vec{u} \]

This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \vec{u} / \partial a_{ij} \)

Either obtain another independent equation or eliminate one of the unknown variables from this equation.
Eigenvalue Problem

- Consider the right eigenvalue problem

\[ \sim \hat{A} \hat{u} = \lambda \hat{u} \]

- Assume that the eigenvalues \( \lambda_k \), for \( k = 1, \ldots, n \) are distinct.
- Hence we have \( n \) linearly independent eigenvectors \( \hat{u}_k \).
- Input parameters \( p \in \{a_{ij}\} \)
- Outputs: \( \lambda_i, \hat{u}_i \)
- FSEs

\[
\sim \frac{\partial \hat{u}}{\partial p} + \frac{\partial A}{\partial p} \hat{u} = \lambda \frac{\partial \hat{u}}{\partial p} + \frac{\partial \lambda}{\partial p} \hat{u}
\]

- This equation has two unknowns \( \partial \lambda / \partial a_{ij} \) and \( \partial \hat{u} / \partial a_{ij} \)
- Either obtain another independent equation or eliminate one of the unknown variables from this equation.

Leon Arriola & James Hyman
Choosing the second strategy, let \( \vec{v} \) be some nonzero, as yet unspecified, vector and take the dot product

\[
\vec{v}^T A \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \sim \frac{\partial \vec{A}}{\partial a_{ij}} \vec{u} = \vec{v}^T \lambda \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \frac{\partial \lambda}{\partial a_{ij}} \vec{u}
\]

Rearranging this equation and writing using the inner product notation \( \langle \vec{a}, \vec{b} \rangle = \vec{b}^T \cdot \vec{a} \), we find

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial \vec{A}}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \left( \sim = A - \lambda \sim \right) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right\rangle
\]

Since \( \left( \sim = A - \lambda I \right)^T = \sim^T - \lambda I \), we can use the Lagrange identity for matrices under the usual inner product

\[
\left\langle \left( \sim = A - \lambda I \right) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right\rangle = \left\langle \frac{\partial \vec{u}}{\partial a_{ij}}, \left( \sim^T - \lambda I \right) \vec{v} \right\rangle
Choosing the second strategy, let $\vec{v}$ be some nonzero, as yet unspecified, vector and take the dot product

$$\vec{v}^T \sim A \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \sim \vec{u} = \vec{v}^T \lambda \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \frac{\partial \lambda}{\partial a_{ij}} \vec{u}$$

Rearranging this equation and writing using the inner product notation $\langle \vec{a}, \vec{b} \rangle = \vec{b}^T \cdot \vec{a}$, we find

$$\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial \sim}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle (A - \lambda \sim) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right\rangle$$

Since $$(A - \lambda \sim)^T = \sim^T - \lambda \sim$$, we can use the Lagrange identity for matrices under the usual inner product

$$\left\langle (A - \lambda \sim) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right\rangle = \left\langle \frac{\partial \vec{u}}{\partial a_{ij}}, (A^T - \lambda \sim) \vec{v} \right\rangle$$
Choosing the second strategy, let \( \vec{v} \) be some nonzero, as yet unspecified, vector and take the dot product

\[
\vec{v}^T A \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \overline{\partial A} \vec{u} = \vec{v}^T \lambda \frac{\partial \vec{u}}{\partial a_{ij}} + \vec{v}^T \overline{\frac{\partial \lambda}{\partial a_{ij}}} \vec{u}
\]

Rearranging this equation and writing using the inner product notation \(<\vec{a}, \vec{b}> = \vec{b}^T \cdot \vec{a}\), we find

\[
\frac{\partial \lambda}{\partial a_{ij}} <\vec{u}, \vec{v}> = \left< \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} > + \left< (A - \lambda I) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right>
\]

Since Because \((A - \lambda I)^T = A^T - \lambda I\), we can use the Lagrange identity for matrices under the usual inner product

\[
\left< (A - \lambda I) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right> = \left< \frac{\partial \vec{u}}{\partial a_{ij}}, (A^T - \lambda I) \vec{v} \right>
\]
Eigenvalue Problem

Compare

\[ \frac{\partial \lambda}{\partial a_{ij}} \langle \tilde{u}, \tilde{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \tilde{u}, \tilde{v} \right\rangle + \left\langle \left( \tilde{A} - \lambda \tilde{I} \right) \frac{\partial \tilde{u}}{\partial a_{ij}}, \tilde{v} \right\rangle \]

with

\[ \frac{\partial \lambda}{\partial a_{ij}} \langle \tilde{u}, \tilde{v} \rangle = \left\langle \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{u}, \tilde{v} \right\rangle + \left\langle \frac{\partial \tilde{u}}{\partial a_{ij}}, \left( \tilde{A}^T - \lambda \tilde{I} \right) \tilde{v} \right\rangle \]

Annihilate the second inner product by forcing the condition

\[ \tilde{A}^T - \lambda \tilde{I} = 0 \]

Adjoint problem is left eigenvalue problem

\[ \tilde{A}^T \tilde{v} = \lambda \tilde{v} \]
Eigenvalue Problem

- Compare

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \left( \sim - \lambda I \right) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \right\rangle
\]

- with

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \frac{\partial \vec{u}}{\partial a_{ij}}, \left( \sim^T - \lambda I \right) \vec{v} \right\rangle
\]

- Annihilate the second inner product by forcing the condition

\[
\sim^T - \lambda I = 0
\]

- Adjoint problem is left eigenvalue problem

\[
\sim^T \vec{v} = \lambda \vec{v}
\]
Eigenvalue Problem

- Compare

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \tilde{u}, \tilde{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \tilde{u}, \tilde{v} \right\rangle + \left\langle \left( \tilde{A} - \lambda \tilde{I} \right) \frac{\partial \tilde{u}}{\partial a_{ij}}, \tilde{v} \right\rangle
\]

- with

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \tilde{u}, \tilde{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \tilde{u}, \tilde{v} \right\rangle + \left\langle \frac{\partial \tilde{u}}{\partial a_{ij}}, \left( \tilde{A}^T - \lambda \tilde{I} \right) \tilde{v} \right\rangle
\]

- Annihilate the second inner product by forcing the condition

\[
\tilde{A}^T - \lambda \tilde{I} = 0
\]

- Adjoint problem is left eigenvalue problem

\[
\tilde{A}^T \tilde{v} = \lambda \tilde{v}
\]
Eigenvalue Problem

- Compare

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \rangle + \langle \left( \sim A - \lambda I \right) \frac{\partial \vec{u}}{\partial a_{ij}}, \vec{v} \rangle
\]

with

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \rangle + \langle \frac{\partial \vec{u}}{\partial a_{ij}}, \left( \sim A^T - \lambda I \right) \vec{v} \rangle
\]

- Annihilate the second inner product by forcing the condition

\[
\sim A^T - \lambda I = 0
\]

- Adjoint problem is left eigenvalue problem

\[
\sim A^T \vec{v} = \lambda \vec{v}
\]
Now

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \frac{\partial \vec{u}}{\partial a_{ij}}, (A^T - \lambda I) \vec{v} \right\rangle
\]

reduces to

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle = u_j v_i
\]

For the \( k \)th right & left eigenvalue problems

\[
\vec{A} \vec{u}_k = \lambda_k \vec{u}_k \quad \text{and} \quad \vec{A}^T \vec{v}_k = \lambda_k \vec{v}_k
\]

It can be shown that

\[
\langle \vec{u}_k, \vec{v}_k \rangle \neq 0
\]
Eigenvalue Problem

Now

\[ \frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \frac{\partial \vec{u}}{\partial a_{ij}}, (\tilde{A}^T - \lambda I) \vec{v} \right\rangle \]

reduces to

\[ \frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle = u_j v_i \]

For the kth right & left eigenvalue problems

\[ \sim \vec{u}_k = \lambda_k \vec{u}_k \quad \text{and} \quad \sim \vec{v}_k = \lambda_k \vec{v}_k \]

It can be shown that

\[ \langle \vec{u}_k, \vec{v}_k \rangle \neq 0 \]
Eigenvalue Problem

Now
\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \frac{\partial u}{\partial a_{ij}}, \left( A^T - \lambda I \right) \vec{v} \right\rangle
\]

reduces to

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle = u_j v_i
\]

For the \( k \)th right & left eigenvalue problems

\( A \vec{u}_k = \lambda_k \vec{u}_k \) and \( A^T \vec{v}_k = \lambda_k \vec{v}_k \)

It can be shown that
\[ \langle \vec{u}_k, \vec{v}_k \rangle \neq 0 \]
Eigenvalue Problem

Now

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle + \left\langle \frac{\partial \vec{u}}{\partial a_{ij}} , (A^T - \lambda I) \vec{v} \right\rangle
\]

reduces to

\[
\frac{\partial \lambda}{\partial a_{ij}} \langle \vec{u}, \vec{v} \rangle = \left\langle \frac{\partial A}{\partial a_{ij}} \vec{u}, \vec{v} \right\rangle = u_j v_i
\]

For the kth right & left eigenvalue problems

\[
\sim \vec{u}_k = \lambda_k \vec{u}_k \quad \text{and} \quad \sim A^T \vec{v}_k = \lambda_k \vec{v}_k
\]

It can be shown that

\[
\langle \vec{u}_k, \vec{v}_k \rangle \neq 0
\]
Eigenvalue Problem

- Right eigenvalue problem (forward problem)

\[ A \tilde{u}_k = \lambda_k \tilde{u}_k \]

- Derivative of the eigenvalue:

\[ \frac{\partial \lambda_k}{\partial a_{ij}} = \frac{(k) u_j (k) v_i}{\langle \tilde{u}_k, \tilde{v}_k \rangle} \]

- Associated left eigenvalue problem (adjoint problem)

\[ A^T \tilde{v}_k = \lambda_k \tilde{v}_k \]
Eigenvalue Problem

- Right eigenvalue problem (forward problem)

\[ A \vec{u}_k = \lambda_k \vec{u}_k \]

- Derivative of the eigenvalue:

\[ \frac{\partial \lambda_k}{\partial a_{ij}} = \frac{(k) u_j (k) v_i}{\langle \vec{u}_k, \vec{v}_k \rangle} \]

- Associated left eigenvalue problem (adjoint problem)

\[ A^T \vec{v}_k = \lambda_k \vec{v}_k \]
Eigenvalue Problem

- Right eigenvalue problem (forward problem)

\[ A\tilde{u}_k = \lambda_k \tilde{u}_k \]

- Derivative of the eigenvalue:

\[
\frac{\partial \lambda_k}{\partial a_{ij}} = \frac{(k)u_j (k)v_i}{\langle \tilde{u}_k, \tilde{v}_k \rangle}
\]

- Associated left eigenvalue problem (adjoint problem)

\[ A^T \tilde{v}_k = \lambda_k \tilde{v}_k \]
Next, we determine $\partial \vec{u} / \partial a_{ij}$.

Normalize the right eigenvectors

$$\langle \vec{u}_k, \vec{u}_k \rangle = 1.$$ 

Fix the indexes $i,j$ and differentiating this condition gives

$$\vec{u}_k^T \frac{\partial \vec{u}_k}{\partial a_{ij}} + \frac{\partial \vec{u}_k^T}{\partial a_{ij}} \vec{u}_k = 0.$$ 

Now use the identity $\vec{a}^T \cdot \vec{b} = \vec{b}^T \cdot \vec{a}$

$$\frac{\partial \vec{u}_k^T}{\partial a_{ij}} \vec{u}_k = \vec{u}_k^T \frac{\partial \vec{u}_k}{\partial a_{ij}},$$

which gives the result that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal, i.e.,

$$\langle \frac{\partial \vec{u}_k}{\partial a_{ij}}, \vec{u}_k \rangle = 0, \quad \text{for } k = 1, \ldots, n.$$
Next, we determine $\partial \vec{u} / \partial a_{ij}$.

Normalize the right eigenvectors

$$\langle \vec{u}_k, \vec{u}_k \rangle = 1.$$  

Fix the indexes $i, j$ and differentiating this condition gives

$$\vec{u}_k^T \frac{\partial \vec{u}_k}{\partial a_{ij}} + \frac{\partial \vec{u}_k^T}{\partial a_{ij}} \vec{u}_k = 0.$$  

Now use the identity $\vec{a}^T \cdot \vec{b} = \vec{b}^T \cdot \vec{a}$

$$\frac{\partial \vec{u}_k^T}{\partial a_{ij}} \vec{u}_k = \vec{u}_k^T \frac{\partial \vec{u}_k}{\partial a_{ij}},$$  

which gives the result that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal, i.e.,

$$\langle \frac{\partial \vec{u}_k}{\partial a_{ij}}, \vec{u}_k \rangle = 0, \quad \text{for} \quad k = 1, \ldots, n.$$
Next, we determine $\partial \mathbf{u} / \partial a_{ij}$.

 Normalize the right eigenvectors

$$\langle \mathbf{u}_k, \mathbf{u}_k \rangle = 1.$$ 

Fix the indexes $i, j$ and differentiating this condition gives

$$\mathbf{u}_k^T \frac{\partial \mathbf{u}_k}{\partial a_{ij}} + \frac{\partial \mathbf{u}_k^T}{\partial a_{ij}} \mathbf{u}_k = 0.$$ 

Now use the identity $\mathbf{a}^T \cdot \mathbf{b} = \mathbf{b}^T \cdot \mathbf{a}$

$$\frac{\partial \mathbf{u}_k^T}{\partial a_{ij}} \mathbf{u}_k = \mathbf{u}_k^T \frac{\partial \mathbf{u}_k}{\partial a_{ij}},$$

which gives the result that $\mathbf{u}_k$ and $\partial \mathbf{u}_k / \partial a_{ij}$ are orthogonal, i.e.,

$$\left\langle \frac{\partial \mathbf{u}_k}{\partial a_{ij}}, \mathbf{u}_k \right\rangle = 0, \quad \text{for} \quad k = 1, \ldots n$$
Next, we determine $\partial \tilde{u} / \partial a_{ij}$.

Normalize the right eigenvectors

$$\langle \tilde{u}_k, \tilde{u}_k \rangle = 1.$$  

Fix the indexes $i,j$ and differentiating this condition gives

$$\tilde{u}_k^T \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial \tilde{u}_k^T}{\partial a_{ij}} \tilde{u}_k = 0.$$  

Now use the identity $\tilde{a}^T \cdot \tilde{b} = \tilde{b}^T \cdot \tilde{a}$

$$\frac{\partial \tilde{u}_k}{\partial a_{ij}} \tilde{u}_k = \tilde{u}_k^T \frac{\partial \tilde{u}_k}{\partial a_{ij}},$$

which gives the result that $\tilde{u}_k$ and $\partial \tilde{u}_k / \partial a_{ij}$ are orthogonal, i.e.,

$$\langle \frac{\partial \tilde{u}_k}{\partial a_{ij}}, \tilde{u}_k \rangle = 0, \quad \text{for} \quad k = 1, \ldots n$$
Next, we determine $\partial \tilde{u} / \partial a_{ij}$.

Normalize the right eigenvectors

$$\langle \tilde{u}_k, \tilde{u}_k \rangle = 1.$$ 

Fix the indexes $i, j$ and differentiating this condition gives

$$\tilde{u}_k^T \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial \tilde{u}_k^T}{\partial a_{ij}} \tilde{u}_k = 0.$$

Now use the identity $\tilde{a}^T \cdot \tilde{b} = \tilde{b}^T \cdot \tilde{a}$

$$\frac{\partial \tilde{u}_k^T}{\partial a_{ij}} \tilde{u}_k = \tilde{u}_k^T \frac{\partial \tilde{u}_k}{\partial a_{ij}},$$

which gives the result that $\tilde{u}_k$ and $\partial \tilde{u}_k / \partial a_{ij}$ are orthogonal, i.e.,

$$\left\langle \frac{\partial \tilde{u}_k}{\partial a_{ij}}, \tilde{u}_k \right\rangle = 0, \quad \text{for} \quad k = 1, \ldots, n.$$
Eigenvalue Problem

FSE:

\[ A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \lambda_k \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial \lambda_k}{\partial a_{ij}} \tilde{u}_k \]

Premultiply by \( \tilde{u}_k^T \) and using the orthogonality condition gives

\[ \tilde{u}_k^T A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \tilde{u}_k^T \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \frac{\partial \lambda_k}{\partial a_{ij}} \]

Using the inner product notation

\[ \left\langle A \frac{\partial \tilde{u}_k}{\partial a_{ij}}, \tilde{u}_k \right\rangle = \frac{\partial \lambda_k}{\partial a_{ij}} - \left\langle \frac{\partial A}{\partial a_{ij}} \tilde{u}_k, \tilde{u}_k \right\rangle, \quad \text{for} \quad k = 1, \ldots, N \]
Eigenvalue Problem

- FSE:

\[ A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \lambda_k \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial \lambda_k}{\partial a_{ij}} \tilde{u}_k \]

- Premultiply by $\tilde{u}_k^T$ and using the orthogonality condition gives

\[ \tilde{u}_k^T A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \tilde{u}_k^T \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \frac{\partial \lambda_k}{\partial a_{ij}} \]

- Using the inner product notation

\[ \langle A \frac{\partial \tilde{u}_k}{\partial a_{ij}}, \tilde{u}_k \rangle = \frac{\partial \lambda_k}{\partial a_{ij}} - \langle \frac{\partial A}{\partial a_{ij}} \tilde{u}_k, \tilde{u}_k \rangle, \quad \text{for} \quad k = 1, \ldots, N \]
Eigenvalue Problem

- **FSE:**
  \[
  A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \lambda_k \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \frac{\partial \lambda_k}{\partial a_{ij}} \tilde{u}_k
  \]

- Premultiply by \(\tilde{u}_k^T\) and using the orthogonality condition gives
  \[
  \tilde{u}_k^T A \frac{\partial \tilde{u}_k}{\partial a_{ij}} + \tilde{u}_k^T \frac{\partial A}{\partial a_{ij}} \tilde{u}_k = \frac{\partial \lambda_k}{\partial a_{ij}}
  \]

- Using the inner product notation
  \[
  \left< A \frac{\partial \tilde{u}_k}{\partial a_{ij}}, \tilde{u}_k \right> = \frac{\partial \lambda_k}{\partial a_{ij}} - \left< \frac{\partial A}{\partial a_{ij}} \tilde{u}_k, \tilde{u}_k \right>, \quad \text{for} \quad k = 1, \ldots, N
  \]
To find an explicit expression for $\frac{\partial \mathbf{u}}{\partial a_{ij}}$, we must introduce additional information.

The key to making further progress is to recall that we have assumed that the $N \times N$ matrix $A$ has $N$ distinct eigenvalues, in which case there exists a complete set of $N$ eigenvectors.

Any vector in $\mathbb{C}^N$ can be expressed as a linear combination of the spanning eigenvectors.

Since $\frac{\partial \mathbf{u}}{\partial a_{ij}}$ is an $N \times 1$ vector, we can write this derivative as a linear combination of the eigenvectors.
To find an explicit expression for $\partial \vec{u} / \partial a_{ij}$, we must introduce additional information.

The key to making further progress is to recall that we have assumed that the $N \times N$ matrix $\sim A$ has $N$ distinct eigenvalues, in which case there exists a complete set of $N$ eigenvectors.

Any vector in $\mathbb{C}^N$ can be expressed as a linear combination of the spanning eigenvectors.

Since $\partial \vec{u} / \partial a_{ij}$ is an $N \times 1$ vector, we can write this derivative as a linear combination of the eigenvectors.
To find an explicit expression for $\partial \vec{u} / \partial a_{ij}$, we must introduce additional information.

The key to making further progress is to recall that we have assumed that the $N \times N$ matrix $A$ has $N$ distinct eigenvalues, in which case there exists a complete set of $N$ eigenvectors.

Any vector in $\mathbb{C}^N$ can be expressed as a linear combination of the spanning eigenvectors.

Since $\partial \vec{u} / \partial a_{ij}$ is an $N \times 1$ vector, we can write this derivative as a linear combination of the eigenvectors.
To find an explicit expression for $\partial \vec{u}/\partial a_{ij}$, we must introduce additional information.

The key to making further progress is to recall that we have assumed that the $N \times N$ matrix $\widetilde{A}$ has $N$ distinct eigenvalues, in which case there exists a complete set of $N$ eigenvectors.

Any vector in $\mathbb{C}^N$ can be expressed as a linear combination of the spanning eigenvectors.

Since $\partial \vec{u}/\partial a_{ij}$ is an $N \times 1$ vector, we can write this derivative as a linear combination of the eigenvectors.
Eigenvalue Problem

- Define the eigenvector matrices $\mathbf{U}$ and $\mathbf{V}$, whose columns are the individual eigenvectors $\mathbf{u}_k$ and $\mathbf{v}_k$ respectively

$$
\mathbf{U} := (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_N) \quad \& \quad \mathbf{V} := (\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_N)
$$

- Let $\mathbf{Λ}$ be the diagonal matrix of eigenvalues $\lambda_k$

$$
\mathbf{Λ} := \begin{pmatrix}
\lambda_1 & & \\
& \lambda_2 & \\
& & \ddots \\
& & & \lambda_N
\end{pmatrix}
$$

- Using this notation, the right and left eigenvalue problems can be written as

$$
\mathbf{AU} = \mathbf{UΛ} \quad \text{and} \quad \mathbf{A}^T \mathbf{V} = \mathbf{VΛ}
$$
Eigenvalue Problem

- Define the eigenvector matrices $\mathbf{U}$ and $\mathbf{V}$, whose columns are the individual eigenvectors $\mathbf{u}_k$ and $\mathbf{v}_k$ respectively

$$\mathbf{U} := (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_N) \quad \& \quad \mathbf{V} := (\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_N)$$

- Let $\mathbf{\Lambda}$ be the diagonal matrix of eigenvalues $\lambda_k$

$$\mathbf{\Lambda} := \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \& \vdots \\ \mathbf{0} & \lambda_N \end{pmatrix}$$

- Using this notation, the right and left eigenvalue problems can be written as

$$\mathbf{A} \mathbf{U} = \mathbf{U} \mathbf{\Lambda} \quad \text{and} \quad \mathbf{A}^T \mathbf{V} = \mathbf{V} \mathbf{\Lambda}$$
Define the eigenvector matrices $\mathbf{U}$ and $\mathbf{V}$, whose columns are the individual eigenvectors $\mathbf{u}_k$ and $\mathbf{v}_k$ respectively

$$\mathbf{U} := (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_N) \ & \mathbf{V} := (\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_N)$$

Let $\mathbf{\Lambda}$ be the diagonal matrix of eigenvalues $\lambda_k$

$$\mathbf{\Lambda} := \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_N \end{pmatrix}$$

Using this notation, the right and left eigenvalue problems can be written as

$$\mathbf{AU} = \mathbf{U}\mathbf{\Lambda} \ \ \text{and} \ \ \mathbf{A}^T\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$$
Eigenvalue Problem

- Earlier we forced the right and left eigenvectors to be normalized, and therefore the matrix eigenvectors satisfy the identity

\[ V^T U = I \]

- The derivative of the matrix of eigenvectors can be written as a linear combination of the eigenspace

\[ \frac{\partial U}{\partial a_{ij}} \sim UC \]

- Where the coefficient matrix is

\[ C := \begin{pmatrix} c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(N)} \\ c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(N)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_N^{(1)} & c_N^{(2)} & c_N^{(3)} & \cdots & c_N^{(N)} \end{pmatrix} \]
Earlier we forced the right and left eigenvectors to be normalized, and therefore the matrix eigenvectors satisfy the identity
\[ \tilde{V}^T \tilde{U} = I \]

The derivative of the matrix of eigenvectors can written as a linear combination of the eigenspace
\[
\frac{\partial \tilde{U}}{\partial a_{ij}} = \tilde{U} \tilde{C}
\]

where the coefficient matrix is
\[
\tilde{C} := \begin{pmatrix}
c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(N)} \\
c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(N)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_N^{(1)} & c_N^{(2)} & c_N^{(3)} & \cdots & c_N^{(N)}
\end{pmatrix}
\]
Eigenvalue Problem

- Earlier we forced the right and left eigenvectors to be normalized, and therefore the matrix eigenvectors satisfy the identity

\[ \tilde{V}^T \tilde{U} = \tilde{I} \]

- The derivative of the matrix of eigenvectors can written as a linear combination of the eigenspace

\[ \frac{\partial \tilde{U}}{\partial a_{ij}} = \tilde{U} \tilde{C} \]

- where the coefficient matrix is

\[ \tilde{C} := \begin{pmatrix}
  c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(N)} \\
  c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(N)} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  c_N^{(1)} & c_N^{(2)} & c_N^{(3)} & \cdots & c_N^{(N)}
\end{pmatrix} \]
For a fixed eigenvector $\tilde{u}^{(k)}$, the derivative can be expanded as the sum

$$\frac{\partial \tilde{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \tilde{u}^{(1)} + \cdots + c_k^{(k)} \tilde{u}^{(k)} + \cdots c_N^{(k)} \tilde{u}^{(N)}$$

Differentiating the right eigenvector matrix equation gives

$$A \sim \frac{\partial \tilde{U}}{\partial a_{ij}} + \sim \frac{\partial A}{\partial a_{ij}} \tilde{U} = \tilde{U} \sim \frac{\partial \Lambda}{\partial a_{ij}} + \sim \frac{\partial \tilde{U}}{\partial a_{ij}} \Lambda$$

Rearranging we get

$$\tilde{U} \left[ \Lambda, \tilde{C} \right] = \tilde{U} \sim \frac{\partial \Lambda}{\partial a_{ij}} - \sim \frac{\partial A}{\partial a_{ij}} \tilde{U}$$

where $[\cdot, \cdot]$ denotes the commutator bracket

$$\left[ \Lambda, \tilde{C} \right] := \Lambda \tilde{C} - \tilde{C} \Lambda.$$
Eigenvalue Problem

- For a fixed eigenvector $\vec{u}^{(k)}$, the derivative can be expanded as the sum

\[
\frac{\partial \vec{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \vec{u}^{(1)} + \cdots + c_k^{(k)} \vec{u}^{(k)} + \cdots c_N^{(k)} \vec{u}^{(N)}
\]

- Differentiating the right eigenvector matrix equation gives

\[
\tilde{A} \frac{\partial \tilde{U}}{\partial a_{ij}} + \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{U} = \tilde{U} \frac{\partial \tilde{\Lambda}}{\partial a_{ij}} + \frac{\partial \tilde{U}}{\partial a_{ij}} \tilde{\Lambda}
\]

- Rearranging we get

\[
\tilde{U} \left[ \tilde{\Lambda}, \tilde{C} \right] = \tilde{U} \frac{\partial \tilde{\Lambda}}{\partial a_{ij}} + \frac{\partial \tilde{U}}{\partial a_{ij}} \tilde{\Lambda}
\]

where $[\cdot, \cdot]$ denotes the commutator bracket

\[
\left[ \tilde{\Lambda}, \tilde{C} \right] := \tilde{\Lambda} \tilde{C} - \tilde{C} \tilde{\Lambda}
\]
Eigenvalue Problem

For a fixed eigenvector $\vec{u}^{(k)}$, the derivative can be expanded as the sum

$$\frac{\partial \vec{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \vec{u}^{(1)} + \cdots + c_k^{(k)} \vec{u}^{(k)} + \cdots c_N^{(k)} \vec{u}^{(N)}$$

Differentiating the right eigenvector matrix equation gives

$$\sim A \frac{\partial \tilde{U}}{\partial a_{ij}} + \sim \frac{\partial A}{\partial a_{ij}} \tilde{U} = \tilde{U} \frac{\partial \Lambda}{\partial a_{ij}} + \frac{\partial \tilde{U}}{\partial a_{ij}} \Lambda$$

Rearranging we get

$$\tilde{U} \left[ \Lambda, \tilde{C} \right] = \tilde{U} \frac{\partial \Lambda}{\partial a_{ij}} - \frac{\partial A}{\partial a_{ij}} \tilde{U}$$

where $[\cdot, \cdot]$ denotes the commutator bracket

$$\left[ \Lambda, \tilde{C} \right] := \Lambda \tilde{C} - \tilde{C} \Lambda.$$
Eigenvalue Problem

Premultiply by the left eigenvector matrix and use the normalization condition, this equation reduces to

\[
[\Lambda, \vec{C}] = \frac{\partial \Lambda}{\partial a_{ij}} - \vec{V}^T \frac{\partial A}{\partial a_{ij}} \vec{U}
\]

Expanding the commutator bracket we find that

\[
[\Lambda, \vec{C}] = \begin{pmatrix}
0 & c_1^{(2)}(\lambda_1 - \lambda_2) & c_1^{(3)}(\lambda_1 - \lambda_3) & \cdots & c_1^{(N)}(\lambda_1 - \lambda_N) \\
c_2^{(1)}(\lambda_2 - \lambda_1) & 0 & c_2^{(3)}(\lambda_2 - \lambda_3) & \cdots & c_2^{(N)}(\lambda_2 - \lambda_N) \\
c_3^{(1)}(\lambda_3 - \lambda_1) & c_3^{(2)}(\lambda_3 - \lambda_2) & 0 & \cdots & c_3^{(N)}(\lambda_3 - \lambda_N) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_N^{(1)}(\lambda_N - \lambda_1) & c_N^{(2)}(\lambda_N - \lambda_2) & c_N^{(3)}(\lambda_N - \lambda_3) & \cdots & 0
\end{pmatrix}
\]

Since the right side is known, and because we assumed that the eigenvalues are distinct, we can solve for the off–diagonal coefficients

\[
c_l^{(m)} = -\frac{1}{\lambda_l - \lambda_m} \left[ \vec{V}^T \frac{\partial A}{\partial a_{ij}} \vec{U} \right]_{lm}
\]

for \( l \neq m \)
Eigenvalue Problem

Premultiply by the left eigenvector matrix and use the normalization condition, this equation reduces to

\[
\begin{bmatrix}
\Lambda, \tilde{C}
\end{bmatrix} = \frac{\partial \Lambda}{\partial a_{ij}} - \tilde{V}^T \frac{\partial \bar{A}}{\partial a_{ij}} \tilde{U}
\]

Expanding the commutator bracket we find that

\[
\begin{bmatrix}
\Lambda, \tilde{C}
\end{bmatrix} = \begin{pmatrix}
0 & c_1^{(2)} (\lambda_1 - \lambda_2) & c_1^{(3)} (\lambda_1 - \lambda_3) & \cdots & c_1^{(N)} (\lambda_1 - \lambda_N) \\
c_2^{(1)} (\lambda_2 - \lambda_1) & 0 & c_2^{(3)} (\lambda_2 - \lambda_3) & \cdots & c_2^{(N)} (\lambda_2 - \lambda_N) \\
c_3^{(1)} (\lambda_3 - \lambda_1) & c_3^{(2)} (\lambda_3 - \lambda_2) & 0 & \cdots & c_3^{(N)} (\lambda_3 - \lambda_N) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
c_N^{(1)} (\lambda_N - \lambda_1) & c_N^{(2)} (\lambda_N - \lambda_2) & c_N^{(3)} (\lambda_N - \lambda_3) & \cdots & 0
\end{pmatrix}
\]

Since the right side is known, and because we assumed that the eigenvalues are distinct, we can solve for the off–diagonal coefficients

\[
c_l^{(m)} = -\frac{1}{\lambda_l - \lambda_m} \left[ \tilde{V}^T \frac{\partial \bar{A}}{\partial a_{ij}} \tilde{U} \right]_{lm}
\]

for \( l \neq m \)
Premultiply by the left eigenvector matrix and use the normalization condition, this equation reduces to

$$\left[ \Lambda, \vec{C} \right] = \frac{\partial \Lambda}{\partial a_{ij}} - \vec{V}^T \frac{\partial A}{\partial a_{ij}} \vec{U}$$

Expanding the commutator bracket we find that

$$\left[ \Lambda, \vec{C} \right] = \begin{pmatrix} 0 & c_1^{(2)} (\lambda_1 - \lambda_2) & c_1^{(3)} (\lambda_1 - \lambda_3) & \cdots & c_1^{(N)} (\lambda_1 - \lambda_N) \\ c_2^{(1)} (\lambda_2 - \lambda_1) & 0 & c_2^{(3)} (\lambda_2 - \lambda_3) & \cdots & c_2^{(N)} (\lambda_2 - \lambda_N) \\ c_3^{(1)} (\lambda_3 - \lambda_1) & c_3^{(2)} (\lambda_3 - \lambda_2) & 0 & \cdots & c_3^{(N)} (\lambda_3 - \lambda_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_N^{(1)} (\lambda_N - \lambda_1) & c_N^{(2)} (\lambda_N - \lambda_2) & c_N^{(3)} (\lambda_N - \lambda_3) & \cdots & 0 \end{pmatrix}$$

Since the right side is known, and because we assumed that the eigenvalues are distinct, we can solve for the off–diagonal coefficients

$$c_l^{(m)} = -\frac{1}{\lambda_l - \lambda_m} \left[ \vec{V}^T \frac{\partial A}{\partial a_{ij}} \vec{U} \right]_{lm}$$  for  $l \neq m$
Use the fact that the eigenvectors form a basis for $\mathbb{C}^N$.

Need to solve for the scalar diagonal coefficients $c_k^{(k)}$.

Using the fact that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal.

We obtain the equation

$$c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots + c_N^{(k)} \langle \vec{u}^{(N)}, \vec{u}^{(k)} \rangle = 0.$$ 

Solve for

The diagonal coefficients in terms of the known off diagonal coefficients are

$$c_k^{(k)} = - \sum_{\substack{i=1 \atop i \neq k}}^{N} c_i^{(k)} \langle \vec{u}^{(i)}, \vec{u}^{(k)} \rangle$$
Use the fact that the eigenvectors form a basis for $\mathbb{C}^N$.

Need to solve for the scalar diagonal coefficients $c_k^{(k)}$

Using the fact that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal.

We obtain the equation

$$c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots c_N^{(k)} \langle \vec{u}^{(N)}, \vec{u}^{(k)} \rangle = 0.$$  

Solve for

The diagonal coefficients in terms of the known off diagonal coefficients are

$$c_k^{(k)} = - \sum_{\substack{i=1 \atop i \neq k}}^{N} c_i^{(k)} \langle \vec{u}^{(i)}, \vec{u}^{(k)} \rangle$$
Eigenvalue Problem

- Use the fact that the eigenvectors form a basis for $\mathbb{C}^N$.
- Need to solve for the scalar diagonal coefficients $c_{k}^{(k)}$
- Using the fact that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal
- We obtain the equation

$$ c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots c_N^{(k)} \langle \vec{u}^{(N)}, \vec{u}^{(k)} \rangle = 0. $$

Solve for

- The diagonal coefficients in terms of the known off diagonal coefficients are

$$ c_k^{(k)} = - \sum_{i=1 \atop i \neq k}^{N} c_i^{(k)} \langle \vec{u}^{(i)}, \vec{u}^{(k)} \rangle $$
Eigenvalue Problem

- Use the fact that the eigenvectors form a basis for $\mathbb{C}^N$.
- Need to solve for the scalar diagonal coefficients $c_k^{(k)}$.
- Using the fact that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal.
- We obtain the equation

\[ c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots c_N^{(k)} \langle \vec{u}^{(N)}, \vec{u}^{(k)} \rangle = 0. \]

Solve for

- The diagonal coefficients in terms of the known off diagonal coefficients are

\[ c_k^{(k)} = - \sum_{\substack{i=1 \\ i \neq k}}^{N} c_i^{(k)} \langle \vec{u}^{(i)}, \vec{u}^{(k)} \rangle \]
Eigenvalue Problem

- Use the fact that the eigenvectors form a basis for $\mathbb{C}^N$.
- Need to solve for the scalar diagonal coefficients $c_k^{(k)}$.
- Using the fact that $\vec{u}_k$ and $\partial \vec{u}_k / \partial a_{ij}$ are orthogonal.
- We obtain the equation

$$c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots c_N^{(k)} \langle \vec{u}^{(N)}, \vec{u}^{(k)} \rangle = 0.$$

Solve for

- The diagonal coefficients in terms of the known off diagonal coefficients are

$$c_k^{(k)} = - \sum_{i=1, i \neq k}^N c_i^{(k)} \langle \vec{u}^{(i)}, \vec{u}^{(k)} \rangle.$$
Summary of the Eigenvalue Problem

- Forward/Adjoint problems $\tilde{A}\tilde{u} = \lambda\tilde{u}$ and $A^T\tilde{v} = \lambda\tilde{v}$
- Derivative of the eigenvalues

$$\frac{\partial \lambda_k}{\partial a_{ij}} = \frac{(k) u_j (k) v_i}{\langle \tilde{u}_k, \tilde{v}_k \rangle}$$
Summary of the Eigenvalue Problem

- Forward/Adjoint problems $A\vec{u} = \lambda \vec{u}$ and $A^T\vec{v} = \lambda \vec{v}$
- Derivative of the eigenvalues

\[
\frac{\partial \lambda_k}{\partial a_{ij}} = \frac{(k) u_j (k) v_i}{\langle \vec{u}_k, \vec{v}_k \rangle}
\]
Summary of the Eigenvalue Problem

- Derivative of the eigenvectors

\[ \frac{\partial \bar{U}}{\partial a_{ij}} = \bar{UC} \]

- where the off–diagonal coefficients are

\[ c_l^{(m)} = -\frac{1}{\lambda_l - \lambda_m} \left[ \bar{V}^T \frac{\partial \bar{A}}{\partial a_{ij}} \bar{U} \right]_{lm} \quad \text{for} \quad l \neq m \]

- and the diagonal coefficients are

\[ c_k^{(k)} = -\sum_{i=1}^{N} \left. c_i^{(k)} \langle \bar{u}^{(i)}, \bar{u}^{(k)} \rangle \right|_{i \neq k} \]
Summary of the Eigenvalue Problem

- Derivative of the eigenvectors

\[ \frac{\partial \tilde{U}}{\partial a_{ij}} \sim \tilde{U} \sim \tilde{C} \]

- where the off–diagonal coefficients are

\[ c_l^{(m)} = -\frac{1}{\lambda_l - \lambda_m} \left[ \tilde{V}^T \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{U} \right]_{lm} \text{ for } l \neq m \]

- and the diagonal coefficients are

\[ c_k^{(k)} = -\sum_{\substack{i=1 \atop i \neq k}}^{N} c_i^{(k)} \langle \tilde{u}^{(i)}, \tilde{u}^{(k)} \rangle \]
Summary of the Eigenvalue Problem

- Derivative of the eigenvectors
  \[
  \frac{\partial \tilde{U}}{\partial a_{ij}} = \tilde{U} \tilde{C}
  \]

  where the off–diagonal coefficients are
  \[
  c_l^{(m)} = - \frac{1}{\lambda_l - \lambda_m} \left[ \tilde{V}^T \frac{\partial A}{\partial a_{ij}} \tilde{U} \right]_{lm}
  \text{for } l \neq m
  \]

- and the diagonal coefficients are
  \[
  c_k^{(k)} = - \sum_{\substack{i=1 \atop i \neq k}}^N c_i^{(k)} \langle \tilde{u}^{(i)}, \tilde{u}^{(k)} \rangle
  \]
To simplify a mathematical model, where numerous categories of variables exist, one would like to be able to identify those variables that can be safely eliminated without affecting the validity of the model.

In order to not inadvertently eliminate significant variables, one must identify groups of variables that are highly correlated & have strongly interacting mechanisms.

Data contains errors or noise.

Need to estimate the uncertainty in the correlation between variables.

Uncertainty in the data creates uncertainty in the correlation estimates and ultimately in the reduced model.
Dimensionality Reduction

- To simplify a mathematical model, where numerous categories of variables exist, one would like to be able to identify those variables that can be safely eliminated without affecting the validity of the model.

- In order to not inadvertently eliminate significant variables, one must identify groups of variables that are highly correlated & have strongly interacting mechanisms.

- Data contains errors or noise.

- Need to estimate the uncertainty in the correlation between variables.

- Uncertainty in the data creates uncertainty in the correlation estimates and ultimately in the reduced model.
Dimensionality Reduction

- To simplify a mathematical model, where numerous categories of variables exist, one would like to be able to identify those variables that can be safely eliminated without affecting the validity of the model.

- In order to not inadvertently eliminate significant variables, one must identify *groups* of variables that are highly correlated & have strongly interacting mechanisms.

- Data contains errors or noise.
  - Need to estimate the uncertainty in the correlation between variables.
  - Uncertainty in the data creates uncertainty in the correlation estimates and ultimately in the reduced model.
To simplify a mathematical model, where numerous categories of variables exist, one would like to be able to identify those variables that can be safely eliminated without affecting the validity of the model.

In order to not inadvertently eliminate significant variables, one must identify groups of variables that are highly correlated & have strongly interacting mechanisms.

Data contains errors or noise.

Need to estimate the uncertainty in the correlation between variables.

Uncertainty in the data creates uncertainty in the correlation estimates and ultimately in the reduced model.
Dimensionality Reduction

- To simplify a mathematical model, where numerous categories of variables exist, one would like to be able to identify those variables that can be safely eliminated without affecting the validity of the model.
- In order to not inadvertently eliminate significant variables, one must identify *groups* of variables that are highly correlated & have strongly interacting mechanisms.
- Data contains errors or noise.
- Need to estimate the uncertainty in the correlation between variables.
- Uncertainty in the data creates uncertainty in the correlation estimates and ultimately in the reduced model.

Leon Arriola & James Hyman
Highly Correlated Data Sets

- Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.
- Suppose that there exists a medication whose sole purpose is to treat this particular disease.
- The number of prescriptions for this medication and the positive blood test results are highly correlated.
- Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.
- The information contained in these two data sets are redundant.
- Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Highly Correlated Data Sets

- Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.

- Suppose that there exists a medication whose sole purpose is to treat this particular disease.

- The number of prescriptions for this medication and the positive blood test results are highly correlated.

- Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.

- The information contained in these two data sets are redundant.

- Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.

Suppose that there exists a medication whose sole purpose is to treat this particular disease.

The number of prescriptions for this medication and the positive blood test results are highly correlated.

Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.

The information contained in these two data sets are redundant.

Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Highly Correlated Data Sets

- Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.
- Suppose that there exists a medication whose sole purpose is to treat this particular disease.
- The number of prescriptions for this medication and the positive blood test results are highly correlated.
- Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.
- The information contained in these two data sets are redundant.
- Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.

Suppose that there exists a medication whose sole purpose is to treat this particular disease.

The number of prescriptions for this medication and the positive blood test results are highly correlated.

Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.

The information contained in these two data sets are redundant.

Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Highly Correlated Data Sets

- Consider an imaginary disease for which a specific blood test can, with absolute certainty, identify whether the patient has or does not have this disease.
- Suppose that there exists a medication whose sole purpose is to treat this particular disease.
- The number of prescriptions for this medication and the positive blood test results are highly correlated.
- Assuming that the examining physician always prescribes this medication the correlation would in fact be 1.0.
- The information contained in these two data sets are redundant.
- Since the two data sets are so highly correlated, a projection from a 2–dimensional parameter space to a 1–dimensional space would be appropriate.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
  - Syndromic surveillance/biosurveillance data of clinical symptoms such as
    - fever
    - number of hospital admissions
    - over-the-counter medication consumption
    - respiratory complaints
    - school or work absences, etc.
  - While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.
  - Noise in the data causes inaccuracy of any specific numerical assessments or predictions.
  - Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.

Syndromic surveillance/biosurveillance data of clinical symptoms such as
- fever
- number of hospital admissions
- over–the–counter medication consumption
- respiratory complaints
- school or work absences, etc.,

While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

Noise in the data causes inaccuracy of any specific numerical assessments or predictions.

Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.
- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.

- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

- While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.

- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.
- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.

- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

- While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.
- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.
- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.

Syndromic surveillance/biosurveillance data of clinical symptoms such as
- fever
- number of hospital admissions
- over-the-counter medication consumption
- respiratory complaints
- school or work absences, etc.,

While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

Noise in the data causes inaccuracy of any specific numerical assessments or predictions.

Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.
- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,
- While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.
- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.
- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Bio–Syndromic Surveillance

- Consider the scenario where public health officials are monitoring a seasonal outbreak of a disease.

- Syndromic surveillance/biosurveillance data of clinical symptoms such as
  - fever
  - number of hospital admissions
  - over–the–counter medication consumption
  - respiratory complaints
  - school or work absences, etc.,

- While this data is readily available, it does not directly provide accurate numerical quantification of the size of the outbreak.

- Noise in the data causes inaccuracy of any specific numerical assessments or predictions.

- Symptoms such as fever and respiratory complaints have different levels of correlation for different diseases.
Principal Component Analysis

- Principal component analysis (PCA) is a powerful method of modern data analysis that provides a systematic way to reduce the dimension of a complex data set to a lower dimension.
- Can reveal hidden simplified structures that would otherwise go unnoticed.
Principal component analysis (PCA) is a powerful method of modern data analysis that provides a systematic way to reduce the dimension of a complex data set to a lower dimension. Can reveal hidden simplified structures that would otherwise go unnoticed.
Consider an $M \times N$ matrix of data measurements $A$ with $M$ data types and $N$ observations of each data type.

Each $M \times 1$ column of $A$ represents the measurement of data at some time $t_n$ for which there are $N$ time samples.
Consider an $M \times N$ matrix of data measurements $A$ with $M$ data types and $N$ observations of each data type.

Each $M \times 1$ column of $A$ represents the measurement of data at some time $t_n$ for which there are $N$ time samples.
Consider an $M \times N$ matrix of data measurements $A$ with $M$ data types and $N$ observations of each data type.

Each $M \times 1$ column of $A$ represents the measurement of data at some time $t_n$ for which there are $N$ time samples.

$$A = \begin{pmatrix} \cdots & a_{1j} = \text{Temperature} & \cdots \\ a_{2j} = \# \text{Ca}^{2+} \text{ in gap junction} \\ \vdots & a_{3j} = \text{Reflectance} & \vdots \\ \vdots \\ \cdots & a_{Mj} = \text{Voltage} & \cdots \end{pmatrix}_{M \times N}$$
Since any $M \times 1$ vector lies in an $M$–dimensional vector space, then there exists an $M$–dimensional orthonormal basis that spans the vector space.

Goal of PCA is to transform the noisy, and possibly redundant data set to a lower dimensional orthonormal basis.

New basis will filter out the noisy data and reveal hidden structures among the data types.
Since any $M \times 1$ vector lies in an $M$–dimensional vector space, then there exists an $M$–dimensional orthonormal basis that spans the vector space.

Goal of PCA is to transform the noisy, and possibly redundant data set to a lower dimensional orthonormal basis.

New basis will filter out the noisy data and reveal hidden structures among the data types.
Since any $M \times 1$ vector lies in an $M$–dimensional vector space, then there exists an $M$–dimensional orthonormal basis that spans the vector space.

Goal of PCA is to transform the noisy, and possibly redundant data set to a lower dimensional orthonormal basis.

New basis will filter out the noisy data and reveal hidden structures among the data types.
Singular Value Decomposition (SVD)

Let $\sim A$ be a real $M \times N$ matrix and let $r$ denote the rank of $\sim A$.
SVD defines a particular factorization as $\sim A = \sim U \sim \Sigma \sim V^T$ where
- $\sim U$ is an $M \times M$ orthogonal matrix i.e., $(\sim U^T \sim U = I_{M \times M})$
- $\sim V$ is an $N \times N$ orthogonal matrix i.e., $(\sim V^T \sim V = I_{N \times N})$
- the $M \times N$ diagonal matrix $\sim \Sigma$ of singular values
  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0; \sigma_{r+1} = \cdots = \sigma_p = 0$ and
  $p := \min(M, N)$. 
Singular Value Decomposition (SVD)

Let $\sim A$ be a real $M \times N$ matrix and let $r$ denote the rank of $\sim A$.

SVD defines a particular factorization as $\sim A = \sim U \sim \Sigma \sim V^T$ where

- $\sim U$ is an $M \times M$ orthogonal matrix i.e., $(\sim U^T \sim U = I_{M \times M})$
- $\sim V$ is an $N \times N$ orthogonal matrix i.e., $(\sim V^T \sim V = I_{N \times N})$
- the $M \times N$ diagonal matrix $\sim \Sigma$ of singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$; $\sigma_{r+1} = \cdots = \sigma_p = 0$ and $p := \min(M, N)$
Singular Value Decomposition (SVD)

- Let $\sim A$ be a real $M \times N$ matrix and let $r$ denote the rank of $\sim A$.
- SVD defines a particular factorization as $\sim A = \sim U \sim \Sigma \sim V^T$ where
  - $\sim U$ is an $M \times M$ orthogonal matrix i.e., $\left(\sim U^T \sim U = \sim I_{M \times M}\right)$
  - $\sim V$ is an $N \times N$ orthogonal matrix i.e., $\left(\sim V^T \sim V = \sim I_{N \times N}\right)$
  - the $M \times N$ diagonal matrix $\sim \Sigma$ of singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$; $\sigma_{r+1} = \cdots = \sigma_p = 0$ and $p := \min(M, N)$
Singular Value Decomposition (SVD)

- Let $A$ be a real $M \times N$ matrix and let $r$ denote the rank of $A$.
- SVD defines a particular factorization as $A = \sim U \sim \Sigma \sim V^T$ where
  - $U$ is an $M \times M$ orthogonal matrix, i.e., $U^T U = I_{M \times M}$
  - $V$ is an $N \times N$ orthogonal matrix, i.e., $V^T V = I_{N \times N}$
  - the $M \times N$ diagonal matrix $\sim \Sigma$ of singular values
    $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$; $\sigma_{r+1} = \cdots = \sigma_p = 0$ and $p := \min(M, N)$
Let \( \sim A \) be a real \( M \times N \) matrix and let \( r \) denote the rank of \( \sim A \).

SVD defines a particular factorization as \( \sim A = \sim U \sim \Sigma \sim V^T \) where

- \( \sim U \) is an \( M \times M \) orthogonal matrix \( \left( \sim U^T \sim U = \sim I_{M \times M} \right) \)
- \( \sim V \) is an \( N \times N \) orthogonal matrix \( \left( \sim V^T \sim V = \sim I_{N \times N} \right) \)
- the \( M \times N \) diagonal matrix \( \sim \Sigma \) of singular values \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0; \sigma_{r+1} = \cdots = \sigma_p = 0 \) and \( p := \min(M, N) \)
Singular Value Decomposition (SVD)

Let $\sim A$ be a real $M \times N$ matrix and let $r$ denote the rank of $\sim A$.

SVD defines a particular factorization as $\sim A = \sim U \sim \Sigma \sim V^T$ where

- $\sim U$ is an $M \times M$ orthogonal matrix i.e., $(\sim U^T \sim U = I_{M \times M})$
- $\sim V$ is an $N \times N$ orthogonal matrix i.e., $(\sim V^T \sim V = I_{N \times N})$
- the $M \times N$ diagonal matrix $\sim \Sigma$ of singular values

$s_1 \geq s_2 \geq \cdots \geq s_r > 0; s_{r+1} = \cdots = s_p = 0$ and $p := \min(M, N)$

$$
\Sigma = \begin{pmatrix}
s_1 & 0 & \cdots & 0 \\
0 & s_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & s_r
\end{pmatrix}
$$
Find the $M$ columns $\vec{u}^{(m)}$ (called the left singular vectors) of $\sim U$, and the $N$ columns $\vec{v}^{(n)}$ (called the right singular vectors) of $\sim V$, where

$$U := (\vec{u}^{(1)} \quad \vec{u}^{(2)} \quad \ldots \quad \vec{u}^{(M)}) , \quad V := (\vec{v}^{(1)} \quad \vec{v}^{(2)} \quad \ldots \quad \vec{v}^{(N)})$$

by solving the singular value problems

$$A \vec{v} = \sigma \vec{u}, \quad \text{and} \quad A^T \vec{u} = \sigma \vec{v}$$

We will first find $\partial \sigma / \partial a_{ij}$
Singular Value Decomposition (SVD)

- Find the $M$ columns $\vec{u}^{(m)}$ (called the left singular vectors) of $\sim U$, and the $N$ columns $\vec{v}^{(n)}$ (called the right singular vectors) of $\sim V$, where

  $$U := (\vec{u}^{(1)} \quad \vec{u}^{(2)} \quad \cdots \quad \vec{u}^{(M)})$$

  $$V := (\vec{v}^{(1)} \quad \vec{v}^{(2)} \quad \cdots \quad \vec{v}^{(N)})$$

- by solving the singular value problems

  $$A\vec{v} = \sigma \vec{u}, \quad \text{and} \quad A^T \vec{u} = \sigma \vec{v}$$

- We will first find $\partial \sigma / \partial a_{ij}$
Find the $M$ columns $\vec{u}^{(m)}$ (called the left singular vectors) of $\sim U$, and the $N$ columns $\vec{v}^{(n)}$ (called the right singular vectors) of $\sim V$, where

$$\sim U := (\vec{u}^{(1)} \vec{u}^{(2)} \ldots \vec{u}^{(M)})\quad \sim V := (\vec{v}^{(1)} \vec{v}^{(2)} \ldots \vec{v}^{(N)})$$

by solving the singular value problems

$$\sim A \vec{v} = \sigma \vec{u}, \quad \text{and} \quad \sim A^T \vec{u} = \sigma \vec{v}$$

We will first find $\partial \sigma / \partial a_{ij}$
Sensitivity of SVD

- Differentiate the singular problems $A\vec{v} = \sigma \vec{u}$ and $A^T\vec{u} = \sigma \vec{v}$ to get the FSEs.
Differentiate the singular problems $A\tilde{v} = \sigma \tilde{u}$ and $A^T\tilde{u} = \sigma \tilde{v}$ to get the FSEs.
Differentiate the singular problems $A\vec{v} = \sigma \vec{u}$ and $A^T \vec{u} = \sigma \vec{v}$ to get the FSEs

$$A \frac{\partial \vec{v}}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \vec{v} = \sigma \frac{\partial \vec{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \vec{u}$$
Sensitivity of SVD

- Differentiate the singular problems $\tilde{A}\tilde{v} = \sigma \tilde{u}$ and $A^T\tilde{u} = \sigma \tilde{v}$ to get the FSEs

$$\tilde{A} \frac{\partial \tilde{v}}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \tilde{v} = \sigma \frac{\partial \tilde{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \tilde{u}$$

$$A^T \frac{\partial \tilde{u}}{\partial a_{ij}} + \frac{\partial A^T}{\partial a_{ij}} \tilde{u} = \sigma \frac{\partial \tilde{v}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \tilde{v}$$
Sensitivity of SVD

- Differentiate the singular problems $A\vec{v} = \sigma \vec{u}$ and $A^T \vec{u} = \sigma \vec{v}$ to get the FSEs

$$A \frac{\partial \vec{v}}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \vec{v} = \sigma \frac{\partial \vec{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \vec{u}$$

$$A^T \frac{\partial \vec{u}}{\partial a_{ij}} + \frac{\partial A^T}{\partial a_{ij}} \vec{u} = \sigma \frac{\partial \vec{v}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \vec{v}$$

- Problem—3 unknowns but only 2 equations
Sensitivity of SVD

- Since $U$ and $V$ are unitary, the associated singular matrices $\tilde{U}$ and $\tilde{V}$ are normalized, i.e., $U^TU = I$ and $V^TV = I$
- In which case $\tilde{u}^T\tilde{u} = 1$ and $\tilde{v}^T\tilde{v} = 1$
- Using this result we find the orthogonality condition

$$\tilde{u}^T \frac{\partial \tilde{u}}{\partial a_{ij}} = 0 \text{ and } \tilde{v}^T \frac{\partial \tilde{v}}{\partial a_{ij}} = 0$$

- Premultiply the FSE $A \frac{\partial \tilde{V}}{\partial a_{ij}} + \tilde{A} \frac{\partial \tilde{V}}{\partial a_{ij}} = \sigma \frac{\partial \tilde{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \tilde{u}$ by $\tilde{u}^T$ and, using the orthogonality & normalizing conditions the FSE reduces to

$$\tilde{u}^T A \frac{\partial \tilde{V}}{\partial a_{ij}} + \tilde{u}^T \tilde{A} \frac{\partial \tilde{V}}{\partial a_{ij}} = \sigma \tilde{u}^T \frac{\partial \tilde{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \tilde{u}^T \tilde{u} = 0$$
Sensitivity of SVD

- Since $\tilde{U}$ and $\tilde{V}$ are unitary, the associated singular matrices $\tilde{U}$ and $\tilde{V}$ are normalized, i.e., $\tilde{U}^T \tilde{U} = I$ and $\tilde{V}^T \tilde{V} = I$
- In which case $\bar{u}^T \bar{u} = 1$ and $\bar{v}^T \bar{v} = 1$
- Using this result we find the orthogonality condition
  \[ \bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} = 0 \quad \text{and} \quad \bar{v}^T \frac{\partial \bar{v}}{\partial a_{ij}} = 0 \]
- Premultiply the FSE $\tilde{A} \frac{\partial \tilde{v}}{\partial a_{ij}} + \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{v} = \sigma \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u}$ by $\bar{u}^T$ and, using the orthogonality & normalizing conditions the FSE reduces to
  \[ \bar{u}^T A \frac{\partial \tilde{v}}{\partial a_{ij}} + \bar{u}^T \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{v} = \sigma \bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u}^T \bar{u} = 1 \]
Sensitivity of SVD

- Since $U$ and $V$ are unitary, the associated singular matrices $\sim U$ and $\sim V$ are normalized, i.e., $U^T U = I$ and $V^T V = I$
- In which case $\bar{u}^T \bar{u} = 1$ and $\bar{v}^T \bar{v} = 1$
- Using this result we find the orthogonality condition

$$\bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} = 0 \quad \text{and} \quad \bar{v}^T \frac{\partial \bar{v}}{\partial a_{ij}} = 0$$

- Premultiply the FSE $A \frac{\partial \bar{v}}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \bar{v} = \sigma \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u}$ by $\bar{u}^T$ and, using the orthogonality & normalizing conditions the FSE reduces to

$$\bar{u}^T A \frac{\partial \bar{v}}{\partial a_{ij}} + \bar{u}^T \frac{\partial A}{\partial a_{ij}} \bar{v} = \sigma \bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u}^T \bar{u} = 0$$
Sensitivity of SVD

Since \( U \) and \( V \) are unitary, the associated singular matrices \( U \) and \( V \) are normalized, i.e., \( U^T U = I \) and \( V^T V = I \)

In which case \( \bar{u}^T \bar{u} = 1 \) and \( \bar{v}^T \bar{v} = 1 \)

Using this result we find the orthogonality condition

\[
\bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} = 0 \quad \text{and} \quad \bar{v}^T \frac{\partial \bar{v}}{\partial a_{ij}} = 0
\]

Premultiply the FSE \( A \frac{\partial \bar{v}}{\partial a_{ij}} + \frac{\partial A}{\partial a_{ij}} \bar{v} = \sigma \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u} \) by \( \bar{u}^T \) and, using the orthogonality & normalizing conditions the FSE reduces to

\[
\bar{u}^T A \frac{\partial \bar{v}}{\partial a_{ij}} + \bar{u}^T \frac{\partial A}{\partial a_{ij}} \bar{v} = \sigma \bar{u}^T \frac{\partial \bar{u}}{\partial a_{ij}} + \frac{\partial \sigma}{\partial a_{ij}} \bar{u}^T \bar{u} = 0
\]
Sensitivity of SVD

- Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$
- Use this result with the orthogonality condition to eliminate the first term to get
Sensitivity of SVD

- Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$
- Use this result with the orthogonality condition to eliminate the first term to get
Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$

Use this result with the orthogonality condition to eliminate the first term to get

$$\frac{\partial \sigma}{\partial a_{ij}} = \tilde{u}^T A \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$
Sensitivity of SVD

- Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$

- Use this result with the orthogonality condition to eliminate the first term to get

$$\frac{\partial \sigma}{\partial a_{ij}} = \tilde{u}^T A \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$

$$= \sigma \tilde{v}^T \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$

$$= 0$$
Sensitivity of SVD

- Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$
- Use this result with the orthogonality condition to eliminate the first term to get

$$\frac{\partial \sigma}{\partial a_{ij}} = \tilde{u}^T A \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{v}$$

$$= \sigma \tilde{v}^T \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{v}$$

$$= 0$$

$$= \tilde{u}^T \frac{\partial \tilde{A}}{\partial a_{ij}} \tilde{v}$$
Sensitivity of SVD

- Rewrite the singular problem $A^T \tilde{u} = \sigma \tilde{v}^T$ as $\tilde{u}^T A = \sigma \tilde{v}^T$
- Use this result with the orthogonality condition to eliminate the first term to get

$$\frac{\partial \sigma}{\partial a_{ij}} = \tilde{u}^T A \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$

$$= \sigma \tilde{v}^T \frac{\partial \tilde{v}}{\partial a_{ij}} + \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$

$$= 0$$

$$= \tilde{u}^T \frac{\partial A}{\partial a_{ij}} \tilde{v}$$

$$\frac{\partial \sigma}{\partial a_{ij}} = u_i v_j$$
Sensitivity of SVD

- Since the derivative of the singular vector is in $\mathbb{R}^M$ it can be written as a linear combination of the singular vectors.
- Define the unknown coefficient matrix as

$$
\sim C := 
\begin{pmatrix}
  c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(M)} \\
  c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(M)} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  c_M^{(1)} & c_M^{(2)} & c_M^{(3)} & \cdots & c_M^{(M)}
\end{pmatrix}
$$

- In which case the derivative of the singular matrix can be written as

$$
\frac{\partial U}{\partial a_{ij}} \sim \sim = U \sim C
$$

- Singular problems can be written in matrix form

$$
\sim A \sim V = \sim U \sim \sim \Sigma \quad \text{and} \quad \sim A^T \sim U = \sim V \sim \sim \Sigma^T
$$
Sensitivity of SVD

Since the derivative of the singular vector is in $\mathbb{R}^M$ it can be written as a linear combination of the singular vectors.

Define the unknown coefficient matrix as

$$C \sim := \begin{pmatrix}
c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(M)} \\
c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(M)} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
c_M^{(1)} & c_M^{(2)} & c_M^{(3)} & \cdots & c_M^{(M)}
\end{pmatrix}$$

In which case the derivative of the singular matrix can be written as

$$\frac{\partial U \sim}{\partial a_{ij}} = U \sim C \sim$$

Singular problems can be written in matrix form

$A \sim V \sim = U \sim \Sigma \sim$ and $A^{T} \sim U \sim = V \sim \Sigma^{T}$
Sensitivity of SVD

Since the derivative of the singular vector is in $\mathbb{R}^M$ it can be written as a linear combination of the singular vectors.

Define the unknown coefficient matrix as

$$
\sim C := \begin{pmatrix} 
    c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(M)} \\
    c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(M)} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    c_M^{(1)} & c_M^{(2)} & c_M^{(3)} & \cdots & c_M^{(M)} 
\end{pmatrix}
$$

In which case the derivative of the singular matrix can be written as

$$
\frac{\partial \sim U}{\partial a_{ij}} = \sim U \sim C
$$

Singular problems can be written in matrix form

$$
\sim A \sim V = \sim U \sim \sim \Sigma \quad \text{and} \quad \sim A^T \sim U = \sim V \sim \sim \Sigma^T
$$

Leon Arriola & James Hyman
Since the derivative of the singular vector is in $\mathbb{R}^M$ it can be written as a linear combination of the singular vectors.

Define the unknown coefficient matrix as

$$
\sim C := \begin{pmatrix}
c_1^{(1)} & c_1^{(2)} & c_1^{(3)} & \cdots & c_1^{(M)} \\
c_2^{(1)} & c_2^{(2)} & c_2^{(3)} & \cdots & c_2^{(M)} \\
\vdots & \vdots & \vdots & \cdots & \vdots \\
c_M^{(1)} & c_M^{(2)} & c_M^{(3)} & \cdots & c_M^{(M)}
\end{pmatrix}
$$

In which case the derivative of the singular matrix can be written as

$$
\frac{\partial \sim U}{\partial a_{ij}} = \sim U \sim C
$$

Singular problems can be written in matrix form

$$
\sim A \sim V = \sim U \sim \Sigma \quad \text{and} \quad \sim A^{T} \sim U = \sim V \sim \Sigma^{T}
$$
Sensitivity of SVD

- Differentiating the singular matrix equation $A^T U = V \Sigma^T$ gives

$$A^T \frac{\partial U}{\partial a_{ij}} + \frac{\partial A^T}{\partial a_{ij}} U = V \frac{\partial \Sigma^T}{\partial a_{ij}} + \frac{\partial V}{\partial a_{ij}} \Sigma^T$$

- Using the fact that $\frac{\partial U}{\partial a_{ij}}$ can be written as a linear combination of the singular vectors $U$ we get

$$A^T U C - \frac{\partial V}{\partial a_{ij}} \Sigma^T = V \frac{\partial \Sigma^T}{\partial a_{ij}} - \frac{\partial A^T}{\partial a_{ij}} U$$
Sensitivity of SVD

- Differentiating the singular matrix equation $A^T \sim U = \sim V \Sigma^T$ gives

$$A^T \frac{\partial U}{\partial a_{ij}} + \frac{\partial A^T}{\partial a_{ij}} U = V \frac{\partial \Sigma^T}{\partial a_{ij}} + \frac{\partial V}{\partial a_{ij}} \Sigma^T$$

- Using the fact that $\frac{\partial U}{\partial a_{ij}}$ can be written as a linear combination of the singular vectors $U$ we get

$$A^T \sim U \sim C - \frac{\partial V}{\partial a_{ij}} \Sigma^T = V \frac{\partial \Sigma^T}{\partial a_{ij}} - \frac{\partial A^T}{\partial a_{ij}} \sim U$$
Sensitivity of SVD

- Differentiate the singular problem \( A \sim V = U \sim \Sigma \) to obtain

\[
A \frac{\partial V}{\partial a_{ij}} \sim = U \frac{\partial \Sigma}{\partial a_{ij}} \sim + U C \Sigma \sim \sim - \frac{\partial A}{\partial a_{ij}} \sim V \sim
\]

- Premultiply \( A^T \sim \sim U \sim C \sim - \frac{\partial V}{\partial a_{ij}} \sim \Sigma^T \sim = V \frac{\partial \Sigma^T}{\partial a_{ij}} \sim - \frac{\partial A^T}{\partial a_{ij}} \sim U \sim \) by matrix \( A \sim \)

\[
AA^T \sim \sim U \sim C \sim - A \frac{\partial V}{\partial a_{ij}} \sim \Sigma^T \sim = AV \frac{\partial \Sigma^T}{\partial a_{ij}} \sim - A \frac{\partial A^T}{\partial a_{ij}} \sim U \sim
\]

\[
AA^T \sim \sim U \sim C \sim - \left( U \frac{\partial \Sigma}{\partial a_{ij}} \sim + U C \Sigma \sim \sim - \frac{\partial A}{\partial a_{ij}} \sim V \sim \right) \sim \Sigma^T \sim = AV \frac{\partial \Sigma^T}{\partial a_{ij}} \sim
\]

- \( - A \frac{\partial A^T}{\partial a_{ij}} \sim U \sim \)
Sensitivity of SVD

- Differentiate the singular problem \( A V = U \Sigma \) to obtain

\[
\frac{\partial V}{\partial a_{ij}} = \frac{\partial \Sigma}{\partial a_{ij}} + U C \Sigma - \frac{\partial A}{\partial a_{ij}} V
\]

- Premultiply \( A^T U C - \frac{\partial V}{\partial a_{ij}} \Sigma^T = V \frac{\partial \Sigma^T}{\partial a_{ij}} - \frac{\partial A^T}{\partial a_{ij}} U \) by matrix \( A \)

\[
A A^T U C - A \frac{\partial V}{\partial a_{ij}} \Sigma^T = A V \frac{\partial \Sigma^T}{\partial a_{ij}} - A \frac{\partial A^T}{\partial a_{ij}} U
\]

\[
A A^T U C - \left( U \frac{\partial \Sigma}{\partial a_{ij}} + U C \Sigma - \frac{\partial A}{\partial a_{ij}} V \right) \Sigma^T = A V \frac{\partial \Sigma^T}{\partial a_{ij}} - A \frac{\partial A^T}{\partial a_{ij}} U
\]
Sensitivity of SVD

Rearranging so as to isolate the expressions containing $UC$, on the left side of the equation, we get

$$AA^TU\sim C - UC\sim \Sigma \Sigma^T = AV\\frac{\partial \Sigma^T}{\partial a_{ij}} - A\\frac{\partial A^T}{\partial a_{ij}} U + U\\frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T - \frac{\partial A}{\partial a_{ij}} V\Sigma^T$$

In order to simplify this result, consider the left side of this equation
Sensitivity of SVD

- Rearranging so as to isolate the expressions containing $UC$, on the left side of the equation, we get

$$AA^T UC - UC\Sigma\Sigma^T = AV \frac{\partial \Sigma^T}{\partial a_{ij}} - A \frac{\partial A^T}{\partial a_{ij}} U + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T - \frac{\partial A}{\partial a_{ij}} V \Sigma^T$$

- In order to simplify this result, consider the left side of this equation
Rearranging so as to isolate the expressions containing $U\Sigma$, on the left side of the equation, we get

$$AA^T U\Sigma - U\Sigma\Sigma\Sigma^T = AV \frac{\partial \Sigma^T}{\partial a_{ij}} - A \frac{\partial A^T}{\partial a_{ij}} U + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T - \frac{\partial A}{\partial a_{ij}} V\Sigma^T$$

In order to simplify this result, consider the left side of this equation

$$AA^T U\Sigma - U\Sigma\Sigma\Sigma^T = AV\Sigma^T C - U\Sigma\Sigma\Sigma^T$$
Rearranging so as to isolate the expressions containing $UC$, on the left side of the equation, we get

$$AA^T UC - U C \Sigma \Sigma^T = \sim A \left( \sim V \frac{\partial \Sigma^T}{\partial a_{ij}} - \sim A \frac{\partial A^T}{\partial a_{ij}} \sim U + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T - \sim A \frac{\partial A}{\partial a_{ij}} V \Sigma^T \right)$$

In order to simplify this result, consider the left side of this equation

$$AA^T U C - U C \Sigma \Sigma^T = \sim A \Sigma^T \sim C - \sim U C \Sigma \Sigma^T$$

$$= \sim U \Sigma \Sigma^T \sim C - \sim U C \Sigma \Sigma^T$$
Rearranging so as to isolate the expressions containing $\widetilde{UC}$, on the left side of the equation, we get

$$AA^T \widetilde{UC} - \widetilde{UC} \Sigma \Sigma^T = AV \frac{\partial \Sigma^T}{\partial a_{ij}} - A \frac{\partial A^T}{\partial a_{ij}} U + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T - \frac{\partial A}{\partial a_{ij}} V \Sigma^T$$

In order to simplify this result, consider the left side of this equation

$$AA^T \widetilde{UC} - \widetilde{UC} \Sigma \Sigma^T = AV \Sigma^T \widetilde{C} - \widetilde{UC} \Sigma \Sigma^T$$

$$= U \Sigma \Sigma^T \widetilde{C} - \widetilde{UC} \Sigma \Sigma^T$$

$$= U \left[ \Sigma \Sigma^T, \widetilde{C} \right]$$
Rearranging so as to isolate the expressions containing $U\sim C$, on the left side of the equation, we get

$$AA^T \sim U\sim C - \sim U\sim C\sim \sim = AV \sim \frac{\partial \sim}{\partial a_{ij}} \sim - A \sim \frac{\partial A^T}{\partial a_{ij}} \sim U + U \sim \frac{\partial \sim}{\partial a_{ij}} \sim \sim T - \sim \frac{\partial \sim}{\partial a_{ij}} \sim V\sim T$$

In order to simplify this result, consider the left side of this equation

$$AA^T \sim U\sim C - \sim U\sim C\sim \sim = AV\sim T \sim C - \sim U\sim C\sim \sim T$$

$$= U\sim T \sim \sim T C - \sim U\sim C\sim \sim T$$

$$= U \sim [\sim \sim T, C]$$

where $[\cdot, \cdot]$ denotes the commutator bracket

$$[\sim \sim T, C] := \sim \sim T C - C\sim \sim T$$
Rewrite the expression

\[ A \sim V \frac{\partial \Sigma^T}{\partial a_{ij}} + U \Sigma \sim T \frac{\partial \Sigma^T}{\partial a_{ij}} = U \Sigma \sim T \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T \]
Sensitivity of SVD

- Rewrite the expression

\[
\begin{align*}
AV \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T & = U \Sigma \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T \\
& = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right]
\end{align*}
\]
Rewrite the expression

\[ AV \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T = U \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] \]

Next rewrite the expression
Sensitivity of SVD

- Rewrite the expression

\[
AV \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T = U \Sigma \frac{\partial \Sigma^T}{\partial a_{ij}} + U \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right]
\]

- Next rewrite the expression

\[
A \frac{\partial A^T}{\partial a_{ij}} U + V \frac{\partial \Sigma^T}{\partial a_{ij}} = A \frac{\partial A^T}{\partial a_{ij}} U + \frac{\partial A}{\partial a_{ij}} A^T U
\]
Sensitivity of SVD

- Rewrite the expression

\[ AV \sim \frac{\partial \Sigma^T}{\partial a_{ij}} + U \sim \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T \]

\[ = U \sim \frac{\partial \Sigma^T}{\partial a_{ij}} + U \sim \frac{\partial \Sigma}{\partial a_{ij}} \Sigma^T \]

\[ = U \sim \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] \]

- Next rewrite the expression

\[ A \sim \frac{\partial A^T}{\partial a_{ij}} U + \frac{\partial A}{\partial a_{ij}} V \Sigma^T \]

\[ = A \sim \frac{\partial A^T}{\partial a_{ij}} U + \frac{\partial A}{\partial a_{ij}} A^T U \]

\[ = \left( \frac{\partial}{\partial a_{ij}} \left[ A A^T \right] \right) U \]
These simplifications give the system of equations in $c_k^{(l)}$:

$$U \left[ \Sigma \Sigma^T, \tilde{C} \right] = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - \left( \frac{\partial}{\partial a_{ij}} \left[ A A^T \right] \right) U$$

Using the unitary condition the commutator bracket simplifies to the final form:

$$\left[ \Sigma \Sigma^T, \tilde{C} \right] = \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - U^T \left( \frac{\partial}{\partial a_{ij}} \left[ A A^T \right] \right) U$$

Expanding the commutator bracket we find that:

$$\left[ \Sigma \Sigma^T, \tilde{C} \right]_{kl} = \begin{cases} 
0 & k = l \text{ or } k \text{ and } l > r \\
 c_k^{(l)} ((\sigma_k)^2 - (\sigma_l)^2) & k, l \leq r \\
 -c_k^{(l)} (\sigma_l)^2 & l \leq r, k \geq r + 1 \\
 c_k^{(l)} (\sigma_k)^2 & k \leq r, l \geq r + 1
\end{cases}$$

We assumed the singular values are distinct, so we can solve for the off–diagonal coefficients.
Sensitivity of SVD

- These simplifications give the system of equations in $c_k^{(l)}$

$$U \left[ \tilde{\Sigma} \tilde{\Sigma}^T, \tilde{C} \right] = U \frac{\partial}{\partial a_{ij}} \left[ \tilde{\Sigma} \tilde{\Sigma}^T \right] - \left( \frac{\partial}{\partial a_{ij}} \left[ \tilde{A} \tilde{A}^T \right] \right) U$$

- Using the unitary condition the commutator bracket simplifies to the final form

$$\left[ \tilde{\Sigma} \tilde{\Sigma}^T, \tilde{C} \right] = \frac{\partial}{\partial a_{ij}} \left[ \tilde{\Sigma} \tilde{\Sigma}^T \right] - U^T \left( \frac{\partial}{\partial a_{ij}} \left[ \tilde{A} \tilde{A}^T \right] \right) U$$

- Expanding the commutator bracket we find that

$$\left[ \tilde{\Sigma} \tilde{\Sigma}^T, \tilde{C} \right]_{kl} = \begin{cases} 
0 & k = l \text{ or } k \text{ and } l > r \\
c_k^{(l)} \left( (\sigma_k)^2 - (\sigma_l)^2 \right) & k, l \leq r \\
-c_k^{(l)} (\sigma_l)^2 & l \leq r, k \geq r + 1 \\
c_k^{(l)} (\sigma_k)^2 & k \leq r, l \geq r + 1
\end{cases}$$

- We assumed the singular values are distinct, so we can solve for the off–diagonal coefficients.
Sensitivity of SVD

These simplifications gives the system of equations in $c_k^{(l)}$

$$U \left[ \Sigma \Sigma^T, C \right] = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - \left( \frac{\partial}{\partial a_{ij}} \left[ AA^T \right] \right) U$$

Using the unitary condition the commutator bracket simplifies to the final form

$$\left[ \Sigma \Sigma^T, C \right] = \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - U^T \left( \frac{\partial}{\partial a_{ij}} \left[ AA^T \right] \right) U$$

Expanding the commutator bracket we find that

$$\left[ \Sigma \Sigma^T, C \right]_{kl} = \begin{cases} 0 & k = l \text{ or } k \text{ and } l > r \\ c_k^{(l)} \left( (\sigma_k)^2 - (\sigma_l)^2 \right) & k, l \leq r \\ -c_k^{(l)} (\sigma_l)^2 & l \leq r, k \geq r + 1 \\ c_k^{(l)} (\sigma_k)^2 & k \leq r, l \geq r + 1 \end{cases}$$

We assumed the singular values are distinct, so we can solve for the off–diagonal coefficients.
Sensitivity of SVD

- These simplifications give the system of equations in \( c_k^{(l)} \)

\[
U \left[ \Sigma \Sigma^T, C \right] = U \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - \left( \frac{\partial}{\partial a_{ij}} \left[ A A^T \right] \right) U
\]

- Using the unitary condition the commutator bracket simplifies to the final form

\[
\left[ \Sigma \Sigma^T, C \right] = \frac{\partial}{\partial a_{ij}} \left[ \Sigma \Sigma^T \right] - U^T \left( \frac{\partial}{\partial a_{ij}} \left[ A A^T \right] \right) U
\]

- Expanding the commutator bracket we find that

\[
\left[ \Sigma \Sigma^T, C \right]_{kl} = \begin{cases} 
0 & k = l \text{ or } k \text{ and } l > r \\
\frac{c_k^{(l)}}{r} ((\sigma_k)^2 - (\sigma_l)^2) & k, l \leq r, l \leq r, k \geq r + 1 \\
-\frac{c_k^{(l)}}{r} (\sigma_l)^2 & k \leq r, l \geq r + 1 \\
\frac{c_k^{(l)}}{r} (\sigma_k)^2 & k \leq r, l \geq r + 1
\end{cases}
\]

- We assumed the singular values are distinct, so we can solve for the off–diagonal coefficients.
The next task is to find the values of the diagonal coefficients. Once again, we make use of the fact that the singular vectors \( \{ \vec{u}^{(k)} \} \) form a basis for \( \mathbb{R}^M \), that is, for a fixed eigenvector \( \vec{u}^{(k)} \), the derivative is expanded as the sum

\[
\frac{\partial \vec{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \vec{u}^{(1)} + \cdots + c_k^{(k)} \vec{u}^{(k)} + \cdots + c_M^{(k)} \vec{u}^{(M)}
\]

Since the derivative of the singular vector is orthogonal to the singular vector we get

\[
c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots + c_M^{(k)} \langle \vec{u}^{(M)}, \vec{u}^{(k)} \rangle = 0
\]

Since the individual singular vectors are orthonormal, the diagonal coefficients are all identically zero. Using similar methods we can find \( \partial \gamma / \partial a_{ij} \).
The next task is to find the values of the diagonal coefficients.

Once again, we make use of the fact that the singular vectors \( \{ \vec{u}^{(k)} \} \) form a basis for \( \mathbb{R}^M \), that is, for a fixed eigenvector \( \vec{u}^{(k)} \), the derivative is expanded as the sum

\[
\frac{\partial \vec{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \vec{u}^{(1)} + \cdots + c_k^{(k)} \vec{u}^{(k)} + \cdots + c_M^{(k)} \vec{u}^{(M)}
\]

Since the derivative of the singular vector is orthogonal to the singular vector we get

\[
c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots + c_M^{(k)} \langle \vec{u}^{(M)}, \vec{u}^{(k)} \rangle = 0
\]

Since the individual singular vectors are orthonormal, the diagonal coefficients are all identically zero.

Using similar methods we can find \( \partial V / \partial a_{ij} \).
The next task is to find the values of the diagonal coefficients.

Once again, we make use of the fact that the singular vectors \( \{ \vec{u}^{(k)} \} \) form a basis for \( \mathbb{R}^M \), that is, for a fixed eigenvector \( \vec{u}^{(k)} \), the derivative is expanded as the sum

\[
\frac{\partial \vec{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \vec{u}^{(1)} + \cdots + c_k^{(k)} \vec{u}^{(k)} + \cdots + c_M^{(k)} \vec{u}^{(M)}
\]

Since the derivative of the singular vector is orthogonal to the singular vector we get

\[
c_1^{(k)} \langle \vec{u}^{(1)}, \vec{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \vec{u}^{(k)}, \vec{u}^{(k)} \rangle + \cdots + c_M^{(k)} \langle \vec{u}^{(M)}, \vec{u}^{(k)} \rangle = 0
\]

Since the individual singular vectors are orthonormal, the diagonal coefficients are all identically zero.

Using similar methods we can find \( \partial \vec{V} / \partial a_{ij} \).
The next task is to find the values of the diagonal coefficients.

Once again, we make use of the fact that the singular vectors \( \{ \bar{u}^{(k)} \} \) form a basis for \( \mathbb{R}^M \), that is, for a fixed eigenvector \( \bar{u}^{(k)} \), the derivative is expanded as the sum

\[
\frac{\partial \bar{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \bar{u}^{(1)} + \cdots + c_k^{(k)} \bar{u}^{(k)} + \cdots + c_M^{(k)} \bar{u}^{(M)}
\]

Since the derivative of the singular vector is orthogonal to the singular vector we get

\[
c_1^{(k)} \langle \bar{u}^{(1)}, \bar{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \bar{u}^{(k)}, \bar{u}^{(k)} \rangle + \cdots + c_M^{(k)} \langle \bar{u}^{(M)}, \bar{u}^{(k)} \rangle = 0
\]

Since the individual singular vectors are orthonormal, the diagonal coefficients are all identically zero.

Using similar methods we can find \( \partial V / \partial a_{ij} \).
Sensitivity of SVD

- The next task is to find the values of the diagonal coefficients.
- Once again, we make use of the fact that the singular vectors \( \{ \bar{u}^{(k)} \} \) form a basis for \( \mathbb{R}^M \), that is, for a fixed eigenvector \( \bar{u}^{(k)} \), the derivative is expanded as the sum

\[
\frac{\partial \bar{u}^{(k)}}{\partial a_{ij}} = c_1^{(k)} \bar{u}^{(1)} + \cdots + c_k^{(k)} \bar{u}^{(k)} + \cdots + c_M^{(k)} \bar{u}^{(M)}
\]

- Since the derivative of the singular vector is orthogonal to the singular vector we get

\[
c_1^{(k)} \langle \bar{u}^{(1)}, \bar{u}^{(k)} \rangle + \cdots + c_k^{(k)} \langle \bar{u}^{(k)}, \bar{u}^{(k)} \rangle + \cdots + c_M^{(k)} \langle \bar{u}^{(M)}, \bar{u}^{(k)} \rangle = 0
\]

- Since the individual singular vectors are orthonormal, the diagonal coefficients are all identically zero.
- Using similar methods we can find \( \partial \tilde{V} / \partial a_{ij} \).
Formality of the Adjoint Method

Problems which are amenable to the adjoint methodology are those that can be expressed in the form

\[ F(u) = f, \]

where \( F \) is a linear/nonlinear operator \( F : X \rightarrow Y \), and \( f \) is the forward forcing function.

The domain and range \( X \) and \( Y \) are assumed to have sufficiently nice topological properties, for example \( X, Y \in \mathbb{H}, \mathbb{S} \).

Associated with the forward problem is the task of determining the sensitivity of some desired response function(al) \( J(u) \).

The adjoint problem and adjoint variable \( v \in X \) arises through the calculation of the Gâteaux derivative:

\[
F'(u)v := \lim_{\epsilon \to 0} \frac{F(u + \epsilon v) - F(u)}{\epsilon}
\]
Formality of the Adjoint Method

- Problems which are amenable to the adjoint methodology are those that can be expressed in the form

\[ F(u) = f, \]

where \( F \) is a linear/nonlinear operator \( F : X \rightarrow Y \), and \( f \) is the forward forcing function.

- The domain and range \( X \) and \( Y \) are assumed to have sufficiently nice topological properties, for example \( X, Y \in \mathbb{H}, \mathbb{S} \).

- Associated with the forward problem is the task of determining the sensitivity of some desired response function(al) \( J(u) \).

- The adjoint problem and adjoint variable \( v \in X \) arises through the calculation of the Gâteaux derivative:

\[ F'(u)v := \lim_{\epsilon \to 0} \frac{F(u + \epsilon v) - F(u)}{\epsilon} \]
Formality of the Adjoint Method

- Problems which are amenable to the adjoint methodology are those that can be expressed in the form

\[ F(u) = f, \]

where \( F \) is a linear/nonlinear operator \( F : X \rightarrow Y \), and \( f \) is the forward forcing function.

- The domain and range \( X \) and \( Y \) are assumed to have sufficiently nice topological properties, for example \( X, Y \in \mathbb{H}, \mathbb{S} \).

- Associated with the forward problem is the task of determining the sensitivity of some desired response function(al) \( J(u) \).

- The adjoint problem and adjoint variable \( v \in X \) arises through the calculation of the Gâteaux derivative:

\[ F'(u)v := \lim_{\varepsilon \rightarrow 0} \frac{F(u + \varepsilon v) - F(u)}{\varepsilon} \]
Formality of the Adjoint Method

Problems which are amenable to the adjoint methodology are those that can be expressed in the form

\[ F(u) = f, \]

where \( F \) is a linear/nonlinear operator \( F : X \to Y \), and \( f \) is the forward forcing function.

The domain and range \( X \) and \( Y \) are assumed to have sufficiently nice topological properties, for example \( X, Y \in \mathbb{H}, \mathbb{S} \).

Associated with the forward problem is the task of determining the sensitivity of some desired response function(al) \( J(u) \).

The adjoint problem and adjoint variable \( v \in X \) arises through the calculation of the Gâteaux derivative:

\[ F'(u)v := \lim_{\epsilon \to 0} \frac{F(u + \epsilon v) - F(u)}{\epsilon} \]
The notation $F'(u)v$ is intended to suggest that the operator $F$ takes the forward variable $u$, and maps it to an operator $F'$, which now depends on both $u$ as well as the adjoint variable $v$.

Formulate an extended representation of the operator $F$ by using the intermediate–value theorem of nonlinear operators permits us to rewrite the forward operator $F$ in extended form:

$$\Phi(u)u = F(u),$$

The residual operator $\Phi$ is defined in integral form

$$\Phi(u) := \int_{\tau=0}^{1} F'(\tau u) d\tau.$$
Formality of the Adjoint Method

- The notation $F'(u)v$ is intended to suggest that the operator $F$ takes the forward variable $u$, and maps it to an operator $F'$, which now depends on both $u$ as well as the adjoint variable $v$.
- Formulate an extended representation of the operator $F$ by using the intermediate–value theorem of nonlinear operators permits us to rewrite the forward operator $F$ in extended form:

$$\Phi(u)u = F(u),$$

- The residual operator $\Phi$ is defined in integral form

$$\Phi(u) := \int_{\tau=0}^{1} F'(\tau u) \, d\tau.$$
Formality of the Adjoint Method

The notation $F'(u)v$ is intended to suggest that the operator $F$ takes the forward variable $u$, and maps it to an operator $F'$, which now depends on both $u$ as well as the adjoint variable $v$.

Formulate an extended representation of the operator $F$ by using the intermediate–value theorem of nonlinear operators permits us to rewrite the forward operator $F$ in extended form:

$$\Phi(u)u = F(u),$$

The residual operator $\Phi$ is defined in integral form

$$\Phi(u) := \int_{\tau=0}^{1} F'(\tau u) \, d\tau.$$
Formality of the Adjoint Method

Given that an appropriate inner product has been defined, consider the adjoint operation

\[ \langle \Phi(u)v, w \rangle = SC1 + \langle v, \Phi^\dagger(u)w \rangle, \]

where SC1 denotes the 1st solvability condition, and \( \Phi^\dagger \) denotes the adjoint operator associated with the forward operator \( \Phi \).

When SC1 = 0, the result is referred to as the Lagrange identity. The associated generalized adjoint problem is defined as

\[ \Phi^\dagger(u)v = g, \]

where the adjoint forcing function \( g \) has not yet been specified.
Formality of the Adjoint Method

- Given that an appropriate inner product has been defined, consider the adjoint operation

\[
\langle \Phi(u)v, w \rangle = SC1 + \langle v, \Phi^\dagger(u)w \rangle,
\]

where SC1 denotes the 1st solvability condition, and \( \Phi^\dagger \) denotes the adjoint operator associated with the forward operator \( \Phi \).

- When \( SC1 = 0 \), the result is referred to as the Lagrange identity. The associated generalized adjoint problem is defined as

\[
\Phi^\dagger(u)v = g,
\]

where the adjoint forcing function \( g \) has not yet been specified.
Formality of the Adjoint Method

- Taking the dot product of the forward problem with the adjoint solution gives
  \[ \langle \Phi(u)u, v \rangle = \langle f, v \rangle, \]
- Taking the dot product of the adjoint problem with the forward solution gives
Formality of the Adjoint Method

- Taking the dot product of the forward problem with the adjoint solution gives
  \[ \langle \Phi(u)u, v \rangle = \langle f, v \rangle, \]
- Taking the dot product of the adjoint problem with the forward solution gives
  \[ \langle \Phi^\dagger(u)v, u \rangle = \langle g, u \rangle \]
Formality of the Adjoint Method

- Taking the dot product of the forward problem with the adjoint solution gives
  \[ \langle \Phi(u)u, v \rangle = \langle f, v \rangle, \]

- Taking the dot product of the adjoint problem with the forward solution gives
  \[ \langle \Phi^\dagger(u)v, u \rangle = \langle g, u \rangle \]

  \[ \langle v, \Phi(u)u \rangle = \langle g, u \rangle \]
Formality of the Adjoint Method

- Taking the dot product of the forward problem with the adjoint solution gives
  \[ \langle \Phi(u)u, v \rangle = \langle f, v \rangle, \]

- Taking the dot product of the adjoint problem with the forward solution gives
  \[ \langle \Phi^\dagger(u)v, u \rangle = \langle g, u \rangle \]
  \[ \langle v, \Phi(u)u \rangle = \langle g, u \rangle \]
  \[ \langle v, f \rangle = \langle g, u \rangle \]
Formality of the Adjoint Method

- Relating the forward and adjoint problems as $\langle g, u \rangle = \langle f, v \rangle$
- The adjoint forcing function $g$ is cleverly chosen so that $\langle g, u \rangle = J(u)$
Formality of the Adjoint Method

- Relating the forward and adjoint problems as $\langle g, u \rangle = \langle f, v \rangle$
- The adjoint forcing function $g$ is cleverly chosen so that $\langle g, u \rangle = J(u)$
Formality of the Adjoint Method

- Relating the forward and adjoint problems as \( \langle g, u \rangle = \langle f, v \rangle \)
- The adjoint forcing function \( g \) is cleverly chosen so that \( \langle g, u \rangle = J(u) \)

\[
\begin{align*}
\text{Forward Problem} & : \Phi(u)u = f \\
\downarrow & \\
\text{Adjoint Product} & : \langle \Phi(u)v, v \rangle = \langle f, v \rangle \\
\downarrow & \\
\text{Adjoint Response} & : J(u) = \langle f, v \rangle \\
\text{Adjoint Problem} & : \Phi^\dagger(u)v = g \\
\downarrow & \\
\text{Forward Product} & : \langle \Phi^\dagger(u)v, u \rangle = \langle g, u \rangle \\
\downarrow & \\
\text{Forward Response} & : J(u) = \langle g, u \rangle
\end{align*}
\]
“Let the buyer beware!”

In order for an adjoint problem to be defined, an associated inner product structure must exist. **No inner product → No adjoint.**

To determine the sensitivity of the associated functional $J = J(u)$, using the adjoint methodology, the functional must be cleverly written in terms of the inner product.

Once an adjoint problem has been defined, if more than one sensitivity is required, (e.g., recall the case of the sensitivity of SVD), additional information must be introduced to make further progress.

SA as discussed here is local in nature. The estimates of derivatives are valid only in some “small” neighborhood of the specified nominal values of the parameters. For a more global approach, uncertainty quantification methodology should be used.
“Let the buyer beware!”

In order for an adjoint problem to be defined, an associated inner product structure must exist. **No inner product ➞ No adjoint.**

To determine the sensitivity of the associated functional \( J = J(u) \), using the adjoint methodology, the functional must be cleverly written in terms of the inner product.

Once an adjoint problem has been defined, if more than one sensitivity is required, (e.g., recall the case of the sensitivity of SVD), additional information must be introduced to make further progress.

SA as discussed here is local in nature. The estimates of derivatives are valid only in some “small” neighborhood of the specified nominal values of the parameters. For a more global approach, uncertainty quantification methodology should be used.
“Let the buyer beware!”

In order for an adjoint problem to be defined, an associated inner product structure must exist. **No inner product → No adjoint.**

To determine the sensitivity of the associated functional $J = J(u)$, using the adjoint methodology, the functional must be cleverly written in terms of the inner product.

Once an adjoint problem has been defined, if more than one sensitivity is required, (e.g., recall the case of the sensitivity of SVD), additional information must be introduced to make further progress.

SA as discussed here is local in nature. The estimates of derivatives are valid only in some “small” neighborhood of the specified nominal values of the parameters. For a more global approach, uncertainty quantification methodology should be used.
“Let the buyer beware!”

In order for an adjoint problem to be defined, an associated inner product structure must exist. No inner product  \( \implies \) No adjoint.

To determine the sensitivity of the associated functional \( J = J(u) \), using the adjoint methodology, the functional must be cleverly written in terms of the inner product.

Once an adjoint problem has been defined, if more than one sensitivity is required, (e.g., recall the case of the sensitivity of SVD), additional information must be introduced to make further progress.

SA as discussed here is local in nature. The estimates of derivatives are valid only in some “small” neighborhood of the specified nominal values of the parameters. For a more global approach, uncertainty quantification methodology should be used.
Caveat Emptor: When the Adjoint Method Fails

- “Let the buyer beware!”
- In order for an adjoint problem to be defined, an associated inner product structure must exist. **No inner product → No adjoint.**
- To determine the sensitivity of the associated functional $J = J(u)$, using the adjoint methodology, the functional must be cleverly written in terms of the inner product.
- Once an adjoint problem has been defined, if more than one sensitivity is required, (e.g., recall the case of the sensitivity of SVD), additional information must be introduced to make further progress.
- SA as discussed here is local in nature. The estimates of derivatives are valid only in some “small” neighborhood of the specified nominal values of the parameters. For a more global approach, uncertainty quantification methodology should be used.
Sensitivity of the Doubling/Tripling Time

Suppose we have an IVP and we are interested in the time it takes for the solution \( u = u(t) \) to double or triple its initial value i.e., \( u(t_D) = 2u_0 \).

For example, we might wish to know how the doubling time for the number of people infected in an epidemic is affected by changes to a specified parameter via \( \partial t_D / \partial p \).

The typical difficulty is that, in general, we do not have the explicit forward solution, in which case explicit expressions for the desired derivatives are not available.

However, numerical values for these derivatives can be calculated through the numerical solution of the forward sensitivity equation(s).
Sensitivity of the Doubling/Tripling Time

- Suppose we have an IVP and we are interested in the time it takes for the solution \( u = u(t) \) to double or triple its initial value, i.e., \( u(t_D) = 2u_0 \).
- For example, we might wish to know how the doubling time for the number of people infected in an epidemic is affected by changes to a specified parameter via \( \frac{\partial t_D}{\partial p} \).
- The typical difficulty is that, in general, we do not have the explicit forward solution, in which case explicit expressions for the desired derivatives are not available.
- However, numerical values for these derivatives can be calculated through the numerical solution of the forward sensitivity equation(s).
Suppose we have an IVP and we are interested in the time it takes for the solution \( u = u(t) \) to double or triple its initial value i.e., \( u(t_D) = 2u_0 \).

For example, we might wish to know how the doubling time for the number of people infected in an epidemic is affected by changes to a specified parameter via \( \partial t_D / \partial p \).

The typical difficulty is that, in general, we do not have the explicit forward solution, in which case explicit expressions for the desired derivatives are not available.

However, numerical values for these derivatives can be calculated through the numerical solution of the forward sensitivity equation(s).
Suppose we have an IVP and we are interested in the time it takes for the solution $u = u(t)$ to double or triple its initial value, i.e., $u(t_D) = 2u_0$.

For example, we might wish to know how the doubling time for the number of people infected in an epidemic is affected by changes to a specified parameter via $\partial t_D / \partial p$.

The typical difficulty is that, in general, we do not have the explicit forward solution, in which case explicit expressions for the desired derivatives are not available.

However, numerical values for these derivatives can be calculated through the numerical solution of the forward sensitivity equation(s).
Suppose we have an IVP and we are interested in the time it takes for the solution \( u = u(t) \) to double or triple its initial value i.e., \( u(t_D) = 2u_0 \).

For example, we might wish to know how the doubling time for the number of people infected in an epidemic is affected by changes to a specified parameter via \( \partial t_D / \partial p \).

The typical difficulty is that, in general, we do not have the explicit forward solution, in which case explicit expressions for the desired derivatives are not available.

However, numerical values for these derivatives can be calculated through the numerical solution of the forward sensitivity equation(s).
Lemma (Sensitivity of time to attain a multiple of the initial condition)

Let \( u = u(t; p, u_0) \) be the solution to the first order IVP

\[
\frac{du}{dt} = f(u, t; p) \quad \text{with} \quad u(0) = u_0,
\]

where \( f \) is differentiable in \( u, t, \) and \( p \). Let \( t_k \) denote the time \( t \) for which \( u \) attains the value \( u(t_k) = ku_0 \), where \( k > 0 \). The derivatives are \( dt_k/du_0 = 0 \) and \( dt_k/dp \) is given by

\[
\frac{dt_k}{dp} = -\frac{\partial u}{\partial p} \bigg|_{t=t_k} \cdot \frac{1}{f(ku_0, t_k; p)}.
\]
Sensitivity of a Critical Point

- Determine which parameter(s), of an IVP modeling the spread of an epidemic has the most effect on the peak of the infection.
- In other words, we want to determine the sensitivity of a critical point, to parameters or initial conditions.

Lemma (Sensitivity of Critical Points)
The derivative \( \frac{dt_{cp}}{dp} \) and \( \frac{dt_{cp}}{du_0} \) is given by

\[
\frac{dt_{cp}}{dp} = \left. \left( \frac{\partial f}{\partial p} + \frac{\partial f}{\partial u} \frac{\partial u}{\partial p} \right) \right|_{t=t_{cp}}
\]

\[
\frac{dt_{cp}}{du_0} = \left. \frac{\partial f}{\partial u} \right|_{t=t_{cp}}
\]
Sensitivity of a Critical Point

- Determine which parameter(s), of an IVP modeling the spread of an epidemic has the most effect on the peak of the infection.
- In other words, we want to determine the sensitivity of a critical point, to parameters or initial conditions.

**Lemma (Sensitivity of Critical Points)**

The derivative $\frac{dt_{cp}}{dp}$ and $\frac{dt_{cp}}{du_0}$ is given by

$$
\frac{dt_{cp}}{dp} = - \left( \frac{\partial f}{\partial p} + \frac{\partial f}{\partial u} \frac{\partial u}{\partial p} \right) \Bigg|_{t=t_{cp}}, \quad \frac{dt_{cp}}{du_0} = - \frac{\partial f}{\partial u} \frac{\partial u}{\partial t_{cp}}.
$$
Sensitivity of a Critical Point

- Determine which parameter(s), of an IVP modeling the spread of an epidemic has the most effect on the peak of the infection.
- In other words, we want to determine the sensitivity of a critical point, to parameters or initial conditions.

Lemma (Sensitivity of Critical Points)

The derivative $\frac{dt_{cp}}{dp}$ and $\frac{dt_{cp}}{du_0}$ is given by

$$\frac{dt_{cp}}{dp} = -\left(\frac{\partial f}{\partial p} + \frac{\partial f}{\partial u} \frac{\partial u}{\partial p}\right) \bigg|_{t=t_{cp}}, \quad \frac{dt_{cp}}{du_0} = -\frac{\partial f}{\partial u} \frac{\partial u}{\partial u_0} \bigg|_{t=t_{cp}}$$
A commonly occurring model in the biological and electrical engineering sciences is the nonlinear system of ODEs

\[
\frac{dx}{dt} = y \quad \text{and} \quad \frac{dy}{dt} = -\omega^2 x + \lambda (1 - x^2) y
\]

which is often referred to as van der Pol’s equations. E.g. Belousov–Zhabotinski reaction, model of oscillatory cardiac pacemaker, coupled oscillators in the small intestine, etc.

Alternatively, this system can be more conveniently written as the single second order ODE

\[
\frac{d^2x}{dt^2} = \lambda(1 - x^2)\frac{dx}{dt} - \omega^2 x
\]

Parametric plot of \((x(t), y(t))\), where \(\omega = \sqrt{2}, \lambda = 1\), IC’s \(x(0) = 0.001, y(0) = 0.0\), and \(t \in [0, 100]\).
Sensitivity of Periodic Solutions to Parameters

- A commonly occurring model in the biological and electrical engineering sciences is the nonlinear system of ODEs

\[
\frac{dx}{dt} = y \quad \text{and} \quad \frac{dy}{dt} = -\omega^2 x + \lambda (1 - x^2) y
\]

which is often referred to as van der Pol’s equations. E.g. Belousov–Zhabotinski reaction, model of oscillatory cardiac pacemaker, coupled oscillators in the small intestine, etc.

- Alternatively, this system can be more conveniently written as the single second order ODE

\[
\frac{d^2x}{dt^2} = \lambda(1 - x^2)\frac{dx}{dt} - \omega^2 x
\]

- Parametric plot of \((x(t), y(t))\), where \(\omega = \sqrt{2}, \lambda = 1, IC’s x(0) = 0.001, y(0) = 0.0, \) and \(t \in [0, 100]\).
A commonly occurring model in the biological and electrical engineering sciences is the nonlinear system of ODEs

\[
\frac{dx}{dt} = y \quad \text{and} \quad \frac{dy}{dt} = -\omega^2 x + \lambda \left(1 - x^2\right) y
\]

which is often referred to as van der Pol’s equations. E.g. Belousov–Zhabotinski reaction, model of oscillatory cardiac pacemaker, coupled oscillators in the small intestine, etc.

Alternatively, this system can be more conveniently written as the single second order ODE

\[
\frac{d^2x}{dt^2} = \lambda(1 - x^2)\frac{dx}{dt} - \omega^2 x
\]

Parametric plot of \((x(t), y(t))\), where \(\omega = \sqrt{2}, \lambda = 1\), IC’s \(x(0) = 0.001, y(0) = 0.0\), and \(t \in [0, 100]\).
The trajectory starts near the origin and it is evident that after a sufficient amount of time has elapsed, the solution is converging to a periodic orbit, or limit cycle.
Consider the IVP where the forward solution $u$ approaches a limit cycle of period $\mathcal{T}$ as $t \to \infty$

$$u(t + \mathcal{T}; u_0, u_0', p) = u(t; u_0, u_0', p), \quad \forall t \in [0, \infty).$$

As is almost aways the case, a closed form of the forward solution is not available, in which case the derivative $\partial \mathcal{T} / \partial p$ can not be explicitly obtained.
Consider the IVP where the forward solution \( u \) approaches a limit cycle of period \( T \) as \( t \to \infty \)

\[
    u(t + T; u_0, u_0', p) = u(t; u_0, u_0', p), \quad \forall t \in [0, \infty).
\]

As is almost always the case, a closed form of the forward solution is not available, in which case the derivative \( \partial T / \partial p \) can not be explicitly obtained.
The derivative $\partial T/\partial p$ is given by the following formula:

\[
\frac{dT}{dp} = \left. \frac{\partial u(t; u_0, u_0', p)}{\partial p} \right|_{s=t+T} - \left. \frac{\partial u(s; u_0, u_0', p)}{\partial p} \right|_{s=t+T} + \frac{\partial u(t; u_0, u_0', p)}{\partial t}.
\]
Sensitivity of Periodic Solutions to Parameters

The derivative $\partial T / \partial p$ is given by the following

**Lemma (Sensitivity of a periodic function)**

Let $u = u(t; u_0, u_0', p)$ be a family of periodic functions with period $T$. The derivative of the period $T$ with respect to the parameter $p$ is given by

$$
\frac{dT}{dp} = \frac{\partial u(t; u_0, u_0', p)}{\partial p} - \frac{\partial u(s; u_0, u_0', p)}{\partial p} \bigg|_{s=t+T}
$$

Leon Arriola & James Hyman
The End

- Thank You!
- Any questions?
Thank You!

Any questions?