

ECSE 343 Numerical Methods in Engineering

Roni Khazaka

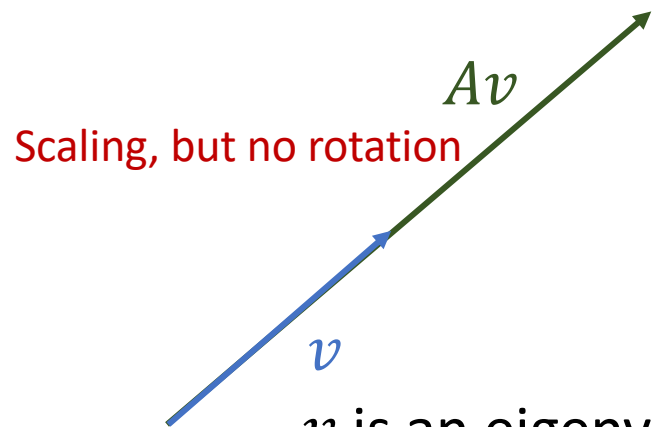
Dept. of Electrical and Computer Engineering

McGill University



McGill

Eigenvalues / Eigenvectors

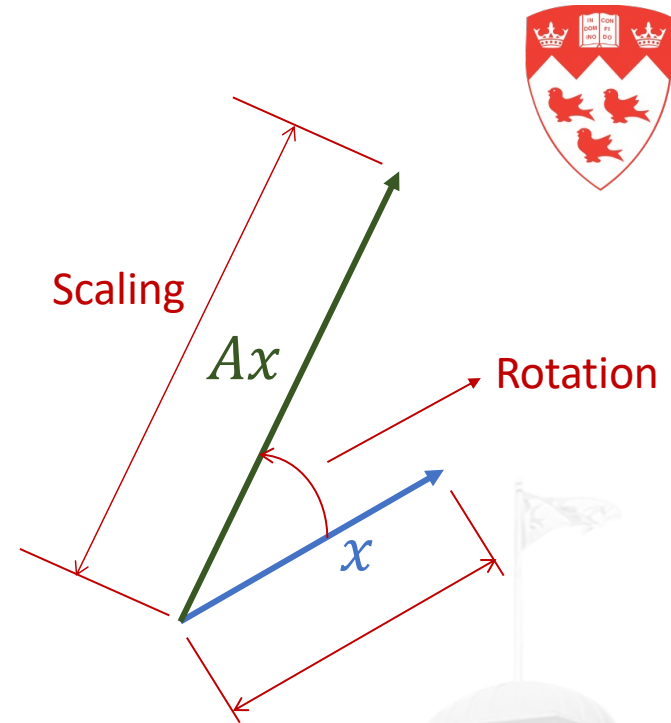


Scaling, but no rotation

v is an eigenvector of A

$$Av = \lambda v$$

eigenvalue (scalar)





Power Method

Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be the eigenvectors of $\mathbf{A} \in \mathbb{R}^{n \times n}$ with corresponding eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_n$. Then:

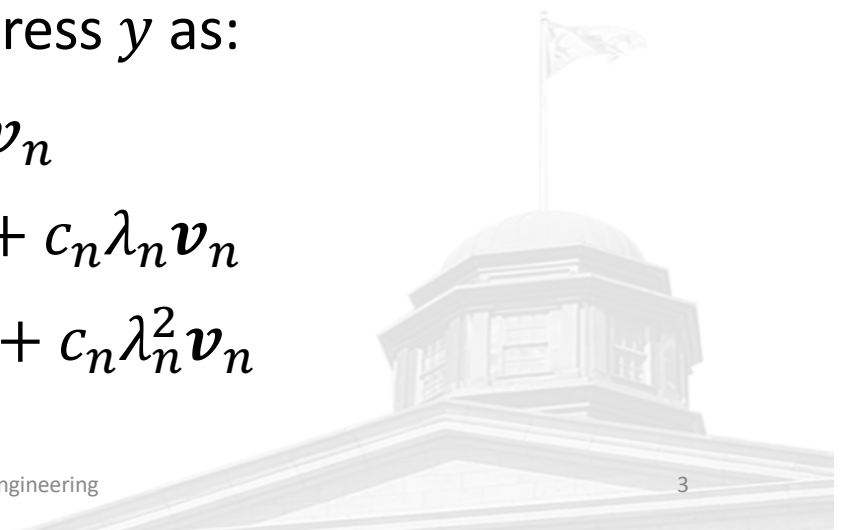
$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Choose a vector $\mathbf{y} \in \mathbb{R}^n$, we can express \mathbf{y} as:

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$$

Then: $\mathbf{A}\mathbf{y} = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \dots + c_n\lambda_n\mathbf{v}_n$

$$\mathbf{A}^2\mathbf{y} = c_1\lambda_1^2\mathbf{v}_1 + c_2\lambda_2^2\mathbf{v}_2 + \dots + c_n\lambda_n^2\mathbf{v}_n$$



Power Method

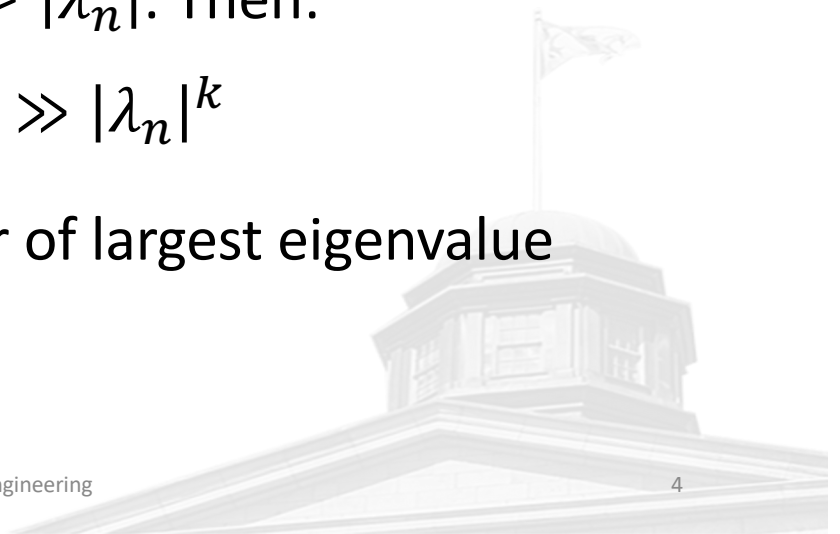


$$\mathbf{A}^k \mathbf{y} = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \cdots + c_n \lambda_n^k \mathbf{v}_n$$

Assume that the eigenvectors and eigen values are ordered such that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. Then:

For very large k : $|\lambda_1|^k \gg |\lambda_2|^k \gg \cdots \gg |\lambda_n|^k$

$$\mathbf{A}^k \mathbf{y} \cong c_1 \lambda_1^k \mathbf{v}_1 \quad \longrightarrow \quad \text{Eigenvector of largest eigenvalue}$$



Power Method (Largest Eigenvalue)



- Choose a vector \mathbf{w}_0
- $\mathbf{w}_1 \leftarrow A\mathbf{w}_0$
- $\mathbf{w}_1 \leftarrow \mathbf{w}_1 / \|\mathbf{w}_1\|$ (Normalize) Length does not matter (avoid large numbers)
- $\mathbf{w}_2 \leftarrow A\mathbf{w}_1$
- $\mathbf{w}_2 \leftarrow \mathbf{w}_2 / \|\mathbf{w}_2\|$ (Normalize)
- $\mathbf{w}_3 \leftarrow A\mathbf{w}_2$
- $\mathbf{w}_3 \leftarrow \mathbf{w}_3 / \|\mathbf{w}_3\|$ (Normalize)
- Continue until $\mathbf{w}_{k+1} \parallel \mathbf{w}_k$





Power Method (Smallest Eigenvalue)

Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be the eigenvectors of $\mathbf{A} \in \mathbb{R}^{n \times n}$ with corresponding eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_n$. Then:

$$\mathbf{A}^{-1}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Choose a vector $\mathbf{y} \in \mathbb{R}^n$, we can express \mathbf{y} as:

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$$

Then: $\mathbf{A}^{-1}\mathbf{y} = c_1\lambda_1^{-1}\mathbf{v}_1 + c_2\lambda_2^{-1}\mathbf{v}_2 + \dots + c_n\lambda_n^{-1}\mathbf{v}_n$

$$\mathbf{A}^{-2}\mathbf{y} = c_1\lambda_1^{-2}\mathbf{v}_1 + c_2\lambda_2^{-2}\mathbf{v}_2 + \dots + c_n\lambda_n^{-2}\mathbf{v}_n$$

Power Method (Smallest Eigenvalue)

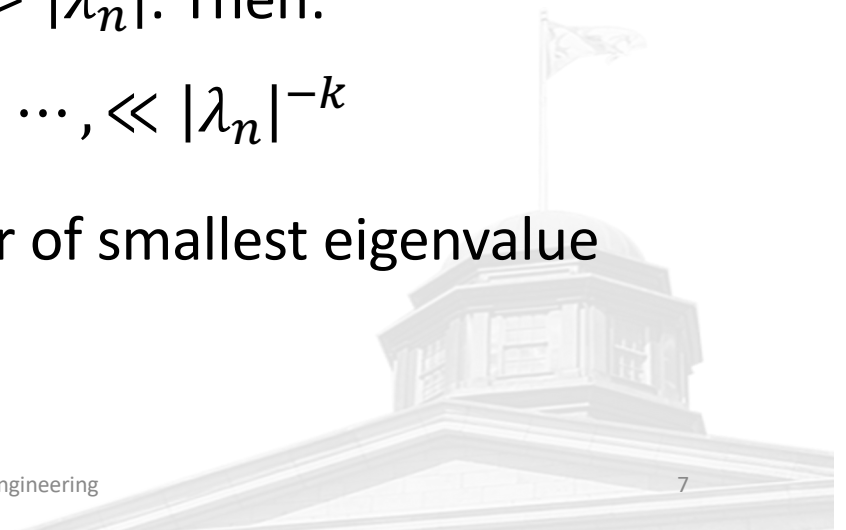


$$\mathbf{A}^{-k} \mathbf{y} = c_1 \lambda_1^{-k} \mathbf{v}_1 + c_2 \lambda_2^{-k} \mathbf{v}_2 + \cdots + c_n \lambda_n^{-k} \mathbf{v}_n$$

Assume that the eigenvectors and eigen values are ordered such that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. Then:

For very large k : $|\lambda_1|^{-k} \ll |\lambda_2|^{-k} \ll \cdots, \ll |\lambda_n|^{-k}$

$\mathbf{A}^{-k} \mathbf{y} \cong c_n \lambda_n^{-k} \mathbf{v}_n \longrightarrow$ Eigenvector of smallest eigenvalue



Power Method (Smallest Eigenvalue)



- Choose a vector \mathbf{w}_0
- Solve: $A\mathbf{w}_1 = \mathbf{w}_0$
- $\mathbf{w}_1 \leftarrow \mathbf{w}_1 / \|\mathbf{w}_1\|$ (Normalize) Length does not matter (avoid large numbers)
- Solve: $A\mathbf{w}_2 = \mathbf{w}_1$
- $\mathbf{w}_2 \leftarrow \mathbf{w}_2 / \|\mathbf{w}_2\|$ (Normalize)
- Solve: $A\mathbf{w}_3 = \mathbf{w}_2$
- $\mathbf{w}_3 \leftarrow \mathbf{w}_3 / \|\mathbf{w}_3\|$ (Normalize)
- Continue until $\mathbf{w}_{k+1} \parallel \mathbf{w}_k$



Second Eigenvalue



Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be the eigenvectors of $\mathbf{A} \in \mathbb{R}^{n \times n}$ with corresponding eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_n$.

Assume that the eigenvectors and eigen values are ordered such that $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$.

Assume that we have already computed \mathbf{v}_1 using the power method, and we would like to compute \mathbf{v}_2





Power Method (2nd Eigenvalue)

Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ be the eigenvectors of $\mathbf{A} \in \mathbb{R}^{n \times n}$ with corresponding eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_n$. Then:

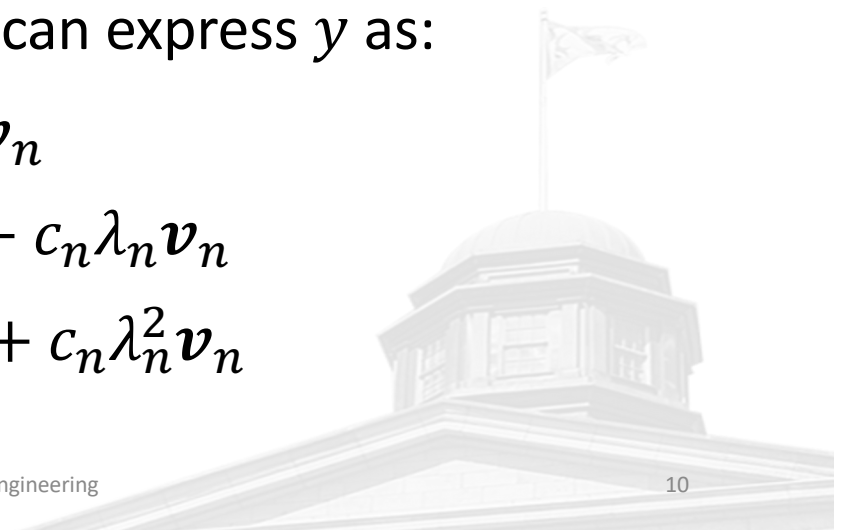
$$\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Choose a vector $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y} \perp \mathbf{v}_1$ we can express \mathbf{y} as:

$$\mathbf{y} = 0\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$$

$$\text{Then: } \mathbf{A}\mathbf{y} = 0\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \dots + c_n\lambda_n\mathbf{v}_n$$

$$\mathbf{A}^2\mathbf{y} = 0\lambda_1^2\mathbf{v}_1 + c_2\lambda_2^2\mathbf{v}_2 + \dots + c_n\lambda_n^2\mathbf{v}_n$$



Power Method

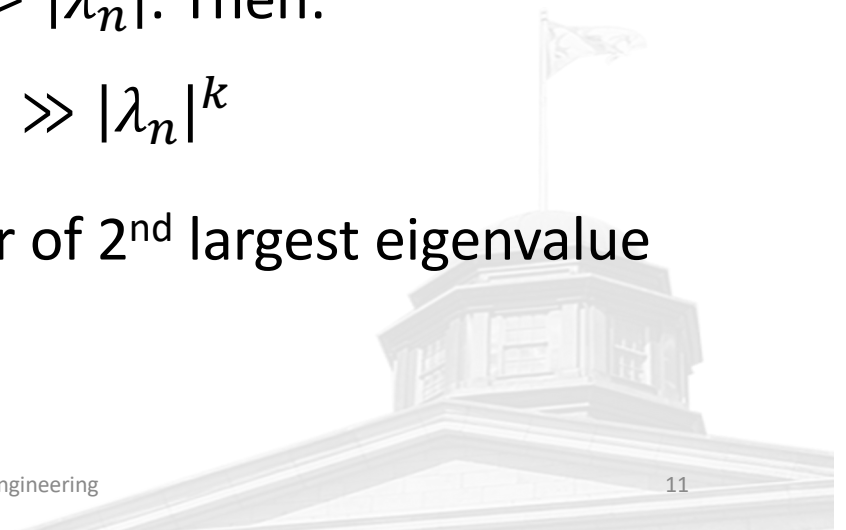


$$\mathbf{A}^k \mathbf{y} = 0\lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \cdots + c_n \lambda_n^k \mathbf{v}_n$$

Assume that the eigenvectors and eigen values are ordered such that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$. Then:

For very large k : $|\lambda_1|^k \gg |\lambda_2|^k \gg \cdots \gg |\lambda_n|^k$

$\mathbf{A}^k \mathbf{y} \cong c_2 \lambda_2^k \mathbf{v}_2 \rightarrow$ Eigenvector of 2nd largest eigenvalue





Power Method (Second Eigenvalue)

- Compute \mathbf{v}_1 (using the power method for example)
- Choose a vector \mathbf{y}_1
- Use QR to obtain: $[\mathbf{v}_1 \quad \mathbf{y}_1] = [q_0 \quad \mathbf{w}_1]R$
 - $\|\mathbf{w}_1\| = 1$ (Normalized)
 - $\mathbf{w}_1 \perp \mathbf{v}_1$
- $\mathbf{y}_2 \leftarrow A\mathbf{w}_1$
- Use QR to obtain: $[\mathbf{v}_1 \quad \mathbf{y}_2] = [q_0 \quad \mathbf{w}_2]R$
 - $\|\mathbf{w}_2\| = 1$ (Normalized)
 - $\mathbf{w}_2 \perp \mathbf{v}_1$
- $\mathbf{y}_3 \leftarrow A\mathbf{w}_2$
- Continue until we converge to an eigenvector.





Some applications of Eigenvalues and Eigenvectors



Least Squares Approximation

Consider n data points (t_i, y_i)

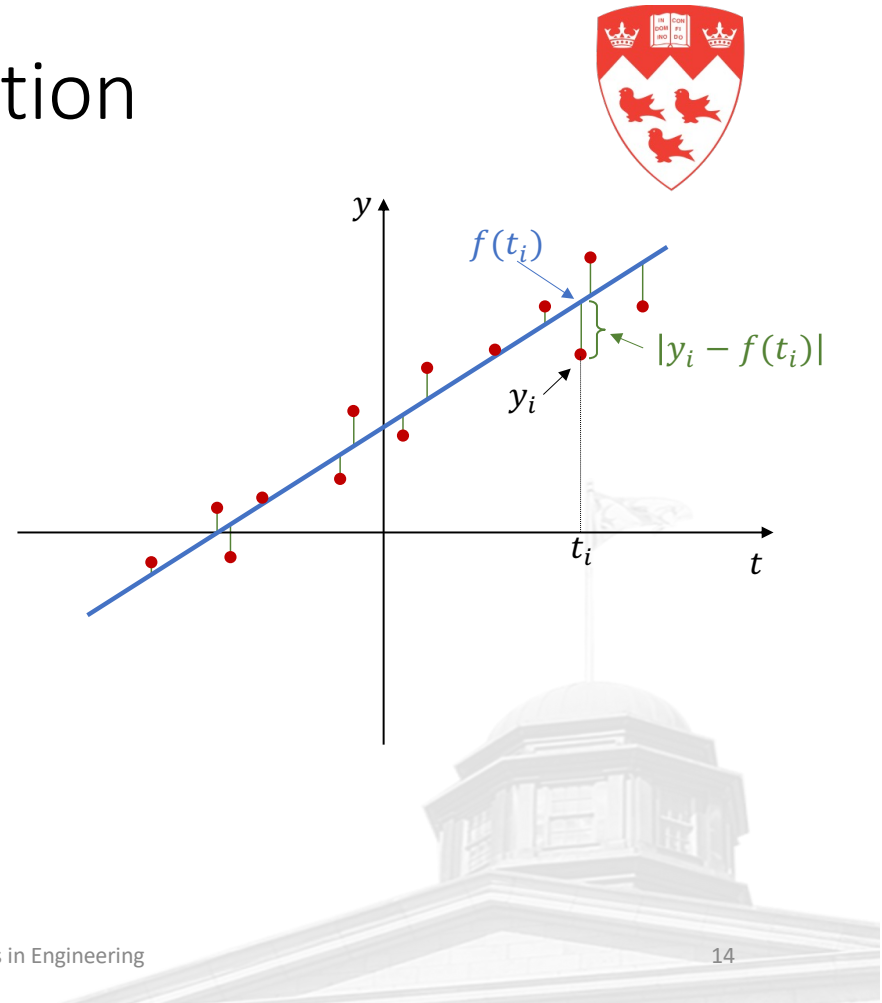
Approximate data with a model:

$$y = f(t) = a_0 + a_1 t$$

a_0 and a_1 are the model parameters.

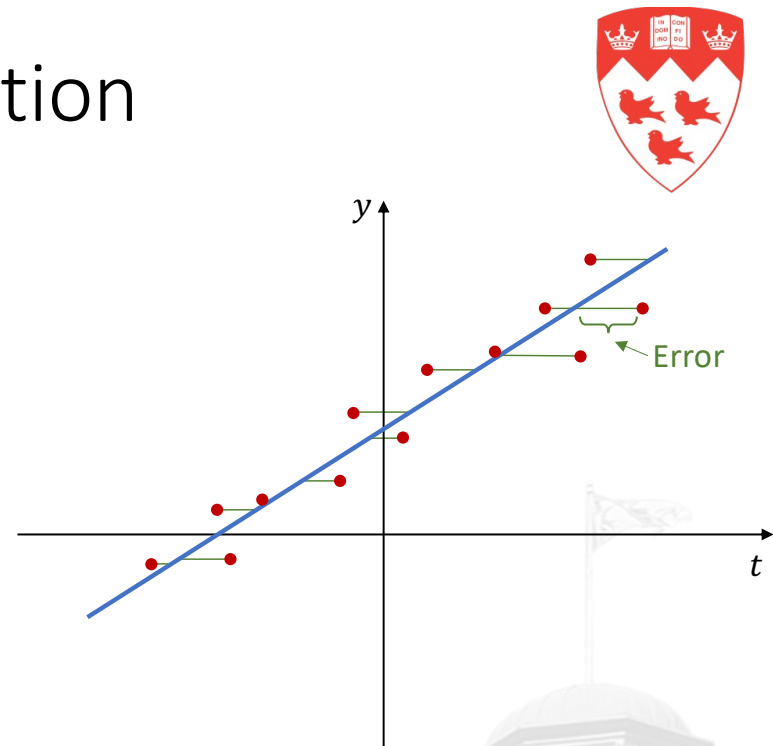
Choose the parameters to minimize:

$$e = \sum_{i=1}^n (f(t_i) - y_i)^2$$



Least Squares Approximation

The problem could have been formulated to minimize the error along the t axis.



Total Least Squares Approximation



Minimize the error in the orthogonal direction.

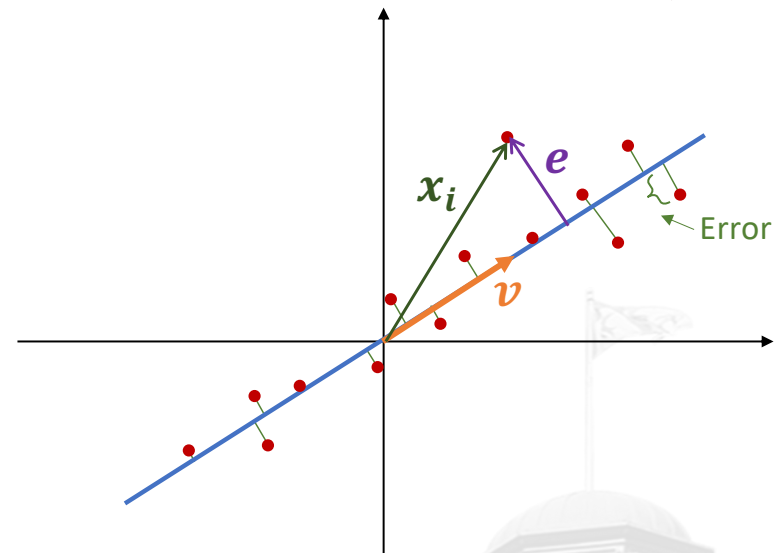
Finding the vector \mathbf{v} defines the line.

Assume \mathbf{v} is normalized: $\|\mathbf{v}\| = 1$

$$\mathbf{e}_i = \mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}$$

Choose \mathbf{v} to minimize:

$$\sum_i \|\mathbf{e}_i\| = \sum_i \|\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}\|$$

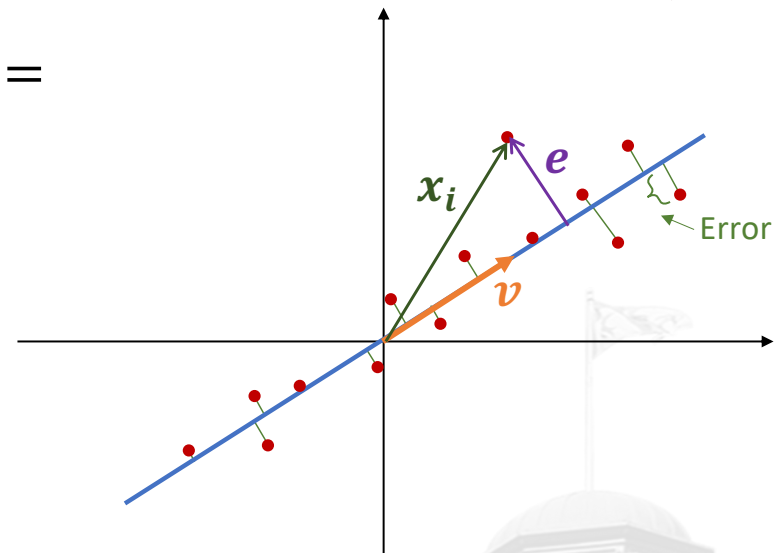




Total Least Squares Approximation

Choose \mathbf{v} to minimize: $\sum_i \|\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}\|^2 =$

$$= \sum_i (\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v})^T (\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v})$$



$$= \sum_i (\mathbf{x}_i^T \mathbf{x}_i - \mathbf{x}_i^T (\mathbf{x}_i^T \mathbf{v}) \mathbf{v} - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}^T \mathbf{x}_i + (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}^T (\mathbf{x}_i^T \mathbf{v}) \mathbf{v})$$



Total Least Squares Approximation

Choose \mathbf{v} to minimize: $\sum_i \|\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}\|^2 = \text{Const.} - \sum_i (\mathbf{x}_i^T \mathbf{v})^2$

$$\mathbf{X} = \left[\begin{bmatrix} \mathbf{x}_1 \end{bmatrix} \quad \begin{bmatrix} \mathbf{x}_2 \end{bmatrix} \quad \cdots \quad \begin{bmatrix} \mathbf{x}_n \end{bmatrix} \right]$$

$$\mathbf{X}^T \mathbf{v} = \begin{bmatrix} \begin{bmatrix} \mathbf{x}_1^T \end{bmatrix} \\ \begin{bmatrix} \mathbf{x}_2^T \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \mathbf{x}_n^T \end{bmatrix} \end{bmatrix} \mathbf{v} = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v} \\ \mathbf{x}_2^T \mathbf{v} \\ \vdots \\ \mathbf{x}_n^T \mathbf{v} \end{bmatrix}$$

$$\sum_i (\mathbf{x}_i^T \mathbf{v})^2 = \|\mathbf{X}^T \mathbf{v}\|_2^2$$

Total Least Squares Approximation



Choose \mathbf{v} to minimize:
$$\sum_i \|\mathbf{x}_i - (\mathbf{x}_i^T \mathbf{v}) \mathbf{v}\|^2 = \text{Const.} - \sum_i (\mathbf{x}_i^T \mathbf{v})^2$$
$$= \text{Const.} - \|\mathbf{X}^T \mathbf{v}\|_2^2$$

Choose \mathbf{v} to maximize: $\|\mathbf{X}^T \mathbf{v}\|_2^2$

Subject to: $\|\mathbf{v}\|_2 = 1$

Choose \mathbf{v} to maximize: $\mathbf{v}^T \mathbf{X} \mathbf{X}^T \mathbf{v}$

Subject to: $\|\mathbf{v}\|_2 = 1$



Total Least Squares Approximation



Choose \mathbf{v} to maximize: $\mathbf{v}^T \mathbf{X} \mathbf{X}^T \mathbf{v}$

Subject to: $\|\mathbf{v}\|_2 = 1$

\mathbf{v} is the eigenvector corresponding to the largest eigenvalue of $\mathbf{X} \mathbf{X}^T$

Proof using the Lagrange Multiplier:

$$\mathcal{L}(\mathbf{v}, \lambda) = \mathbf{v}^T \mathbf{X} \mathbf{X}^T \mathbf{v} - \lambda(\mathbf{v}^T \mathbf{v} - 1)$$

$$\frac{d}{d\mathbf{v}} \mathcal{L}(\mathbf{v}, \lambda) = 2\mathbf{X} \mathbf{X}^T \mathbf{v} - 2\lambda \mathbf{v} = \mathbf{0}$$



Total Least Squares Approximation



Proof using the Lagrange Multiplier:

$$\mathcal{L}(\mathbf{v}, \lambda) = \mathbf{v}^T \mathbf{X} \mathbf{X}^T \mathbf{v} - \lambda(\mathbf{v}^T \mathbf{v} - 1)$$

$$\frac{d}{d\mathbf{v}} \mathcal{L}(\mathbf{v}, \lambda) = 2\mathbf{X} \mathbf{X}^T \mathbf{v} - 2\lambda \mathbf{v} = \mathbf{0}$$

Critical points when \mathbf{v} is an eigenvector: $\mathbf{X} \mathbf{X}^T \mathbf{v} = \lambda \mathbf{v}$

At the critical points: $\mathbf{v}^T \mathbf{X} \mathbf{X}^T \mathbf{v} = \lambda \mathbf{v}^T \mathbf{v} = \lambda$

The maximum occurs when \mathbf{v} is the eigenvector corresponding to the largest eigenvalue of $\mathbf{X} \mathbf{X}^T$



Diagonalization



Consider a full rank matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with distinct eigen values λ_i and corresponding eigenvectors \mathbf{v}_i , $1 \leq i \leq n$.

$$\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \quad \mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2 \quad \cdots \quad \mathbf{A}\mathbf{v}_n = \lambda_n\mathbf{v}_n$$

$$\mathbf{A}[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] = [\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n] = \mathbf{V}\mathbf{\Gamma}$$

$$\begin{aligned} \mathbf{A}\mathbf{V} &= \mathbf{V}\mathbf{\Gamma} & \mathbf{V} &= [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] & \mathbf{\Gamma} &= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \\ \mathbf{V}^{-1}\mathbf{A}\mathbf{V} &= \mathbf{\Gamma} & \mathbf{A} &= \mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1} \end{aligned}$$

Application: Powers of a matrix **A**



$$\mathbf{A} = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1}$$

$$\mathbf{A}^2 = (\mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1})(\mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1}) = \mathbf{V}\mathbf{\Gamma}(\mathbf{V}^{-1}\mathbf{V})\mathbf{\Gamma}\mathbf{V}^{-1} = \mathbf{V}\mathbf{\Gamma}^2\mathbf{V}^{-1}$$

$$\mathbf{A}^m = \mathbf{V}\mathbf{\Gamma}^m\mathbf{V}^{-1}$$

$$\mathbf{\Gamma}^m = \begin{bmatrix} \lambda_1^m & & & \\ & \lambda_1^m & & \\ & & \ddots & \\ & & & \lambda_n^m \end{bmatrix}$$

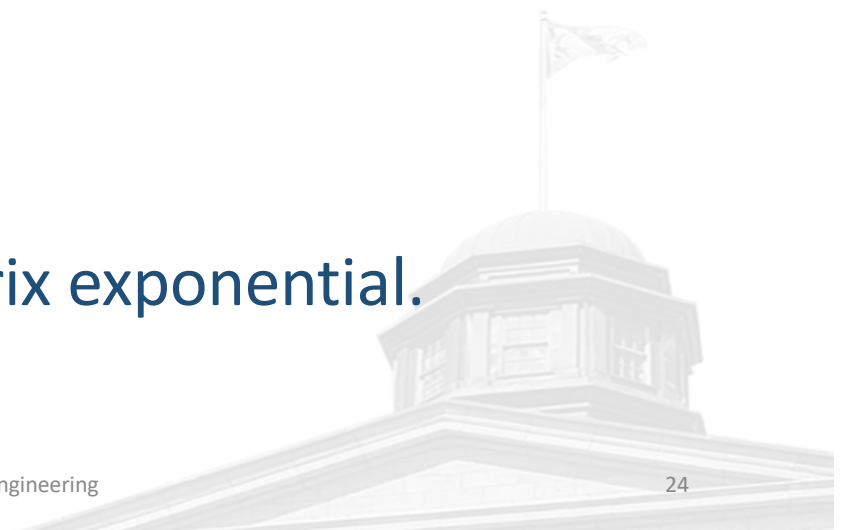


Matlab exp vs expm

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\exp(\mathbf{A}) = \begin{bmatrix} e^{a_{11}} & e^{a_{12}} \\ e^{a_{21}} & e^{a_{22}} \end{bmatrix} \neq e^{\mathbf{A}t}$$

Useful function but not the matrix exponential.



Matlab exp vs expm

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\text{expm}(\mathbf{A}) = \sum_{m=0}^{\infty} \frac{\mathbf{A}^m}{m!} = e^{\mathbf{A}}$$



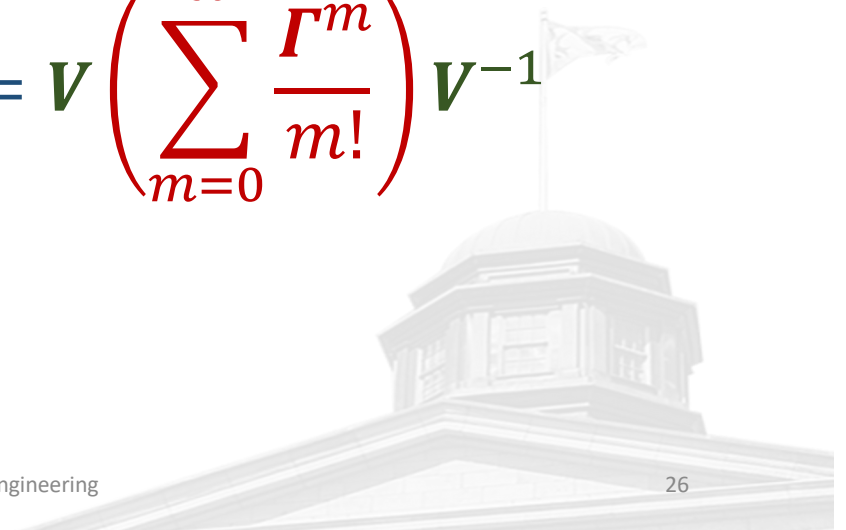


Matlab exp vs expm

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1}$$

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = \sum_{m=0}^{\infty} \frac{\mathbf{V}\mathbf{\Gamma}^m\mathbf{V}^{-1}}{m!} = \mathbf{V} \left(\sum_{m=0}^{\infty} \frac{\mathbf{\Gamma}^m}{m!} \right) \mathbf{V}^{-1}$$

$$e^A = \mathbf{V}(e^{\mathbf{\Gamma}})\mathbf{V}^{-1}$$



Matlab exp vs expm

$$A = V \Gamma V^{-1}$$

$$e^A = V \left(\sum_{m=0}^{\infty} \frac{\Gamma^m}{m!} \right) V^{-1}$$

$$\Gamma^m = \begin{bmatrix} \lambda_1^m & & & \\ & \lambda_1^m & & \\ & & \ddots & \\ & & & \lambda_n^m \end{bmatrix}$$

$$\sum_{m=0}^{\infty} \frac{\Gamma^m}{m!} = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{\lambda_1^m}{m!} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sum_{m=0}^{\infty} \frac{\lambda_n^m}{m!} \end{bmatrix}$$



Matlab exp vs expm

$$A = V\Gamma V^{-1}$$

$$e^A = V \left(\sum_{m=0}^{\infty} \frac{\Gamma^m}{m!} \right) V^{-1}$$

$$\sum_{m=0}^{\infty} \frac{\Gamma^m}{m!} = \begin{bmatrix} e^{\lambda_1^m} & & & \\ & e^{\lambda_2^m} & & \\ & & \ddots & \\ & & & e^{\lambda_n^m} \end{bmatrix} = e^{\Gamma}$$

$$\Gamma^m = \begin{bmatrix} \lambda_1^m & & & \\ & \lambda_1^m & & \\ & & \ddots & \\ & & & \lambda_n^m \end{bmatrix}$$





Decoupled Diff Equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{B}$$

Constant vector

Matrix Exponential

Initial Value Problem:

We know the boundary condition $\mathbf{x}(0)$





Decoupled Diff Equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{B}$$

Constant vector

Matrix Exponential

$$\mathbf{x}(0) = e^{0t} \mathbf{B} = \mathbf{B}$$

$$\mathbf{x}(t) = e^{\mathbf{A}t} \mathbf{x}(0)$$



Initial Value

$$A = V\Gamma V^{-1}$$

$$V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$$

$$\mathbf{x}(0) = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n = V\mathbf{c}$$

$$V\mathbf{c} = \mathbf{x}(0)$$

$$\mathbf{c} = V^{-1}\mathbf{x}(0)$$



$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$



Decoupled Diff Equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t)$$

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) = (\mathbf{V}(e^{\mathbf{\Lambda}t})\mathbf{V}^{-1})(\mathbf{V}\mathbf{c}) = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{c}$$

$$\mathbf{x}(t) = \mathbf{V} \begin{bmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix} = c_1 e^{\lambda_1 t} \mathbf{v}_1 + c_2 e^{\lambda_2 t} \mathbf{v}_2 + \cdots + c_n e^{\lambda_n t} \mathbf{v}_n$$

