Ch 1. Applied Linear Algebra

Yoon Mo Jung

Computational Science and Engineering
Yonsei University

§0 Prelude: Structure of Computational Science and Engineering

• Overview of Computational Science and Engineering

laws rules constraints	$\xrightarrow{\text{modeling}}$	Mathematical problem: equations	$\xrightarrow{\text{Numerical procedures}}$	Linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$

in engineering and physics

Remark *1. Designing and validating numerical procedures is called "Numerical Analysis".*

2. How to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ *is called Numerical Linear Algebra, which is the heart of the Scientific Computing. Strang said* "**Its importance is now recognized.**"

• Four Simplifications

1. **Nonlinear becomes linear.**

   ex) Bending of beam

$$\frac{u''}{(1 + (u')^2)^{3/2}} \approx u'' \quad \text{if } u' \text{ is small.}$$

2. **Continuous becomes discrete.**

$$\text{ex) } f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \approx \frac{f(x_{n+1}) - f(x_n)}{\Delta x},$$

   if Δx is small.

3. **Multidimensional becomes one-dimensional.**

   ex) $u_t = u_{xx}$ (heat equation)

Let $u(t, x) = T(t)X(x)$: separation of variables

4. **Variable coefficients becomes constant coefficients.**

   ex) inhomogeneous heat equation: $u_t = (c(x)u_x)_x$

   with conductivity $c(x)$.

   If $c(x) \approx c$, one may use Fourier transform or FFT.

§1 Applied Linear Algebra

§1.1 Four Special Matrices

- **Two important problems**

1. Solving linear systems: $\boldsymbol{Ax} = \boldsymbol{b}$

   $\boldsymbol{x}$: cause or input, $\boldsymbol{b}$: result or output.

   ex) $\boldsymbol{x}$: displacements, pressures, voltages, concentrations, ...

2. Eigenvalue problem: $\boldsymbol{Ax} = \boldsymbol{\lambda x}$

   eigen means prime.

- **Four special matrices: $\boldsymbol{K_n}$, $\boldsymbol{C_n}$, $\boldsymbol{T_n}$, $\boldsymbol{B_n}$**

$$\boldsymbol{K}_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \boldsymbol{K}_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \boldsymbol{K}_4 = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

Remark *(Good) Properties*

1. *Symmetric ($\boldsymbol{K}_{ij} = \boldsymbol{K}_{ji}$ or $\boldsymbol{K} = \boldsymbol{K}^T$).*

2. *Sparse (lots of zeros if $\boldsymbol{n} \gg 1$).*

3. *Tridiagonal, banded.*

4. *Constant diagonals: called Toeplitz matrix.*

   - Something is not changing when we move in space or time. Shift-invariant or time-invariant.

5. *Invertible ($\exists \boldsymbol{K}^{-1}$ s.t. $\boldsymbol{KK}^{-1} = \boldsymbol{K}^{-1}\boldsymbol{K} = \boldsymbol{I}$).*

   -$\boldsymbol{K}^{-1}$ also symmetric but full.

   - Important: We don't want or need $\boldsymbol{K}^{-1}$ to find $\boldsymbol{u} = \boldsymbol{K}^{-1}\boldsymbol{f}$. All we compute is the solution x.

6. *Positive definite* ($x^T Ax > 0$ *if* $x \neq 0$).

Circulant matrix $$C_4 = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

Remark 1. *Singular, not invertible* ($C_4[1,1,1,1]^T = 0$).

2. *Positive semidefinite* ($x^T Ax \geq 0$).

$$T_2 = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}, \quad T_3 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Gaussian elimination:

$$T = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow{\text{Step 1}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow{\text{Step 2}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = U$$

$$U^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

: The inverse of "difference matrix" is a "sum matrix".

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u_1 - u_2 \\ u_2 - u_3 \\ u_3 - 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 - u_2 \\ u_2 - u_3 \\ u_3 - 0 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Remark: The inverse of triangular matrix is also triangular.

$$\boldsymbol{B}_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad \boldsymbol{B}_3 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

It is positive semidefinite.

$$\boldsymbol{B} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow[\text{Step 1}]{} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \xrightarrow[\text{Step 2}]{} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \boldsymbol{U}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- **Summary**

1. $\boldsymbol{K}_n$ and $\boldsymbol{T}_n$ are invertible and positive definite.
2. $\boldsymbol{C}_n$ and $\boldsymbol{B}_n$ are singular and positive semidefinite.

   The nullspace(kernel) is the constant vector $\boldsymbol{u} = [\boldsymbol{c}, \boldsymbol{c}, \cdots, \boldsymbol{c}]$.

Remark: $\boldsymbol{Bu} = \boldsymbol{f}$ is solvable when $\boldsymbol{f}$ is perpendicular to $\boldsymbol{e} = [1, 1, \cdots, 1]$.

$$\boldsymbol{f} = \boldsymbol{Bu} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} u_1 - u_2 \\ -u_1 2u_2 - u_3 \\ -u_2 + 2u_3 - u_4 \\ -u_3 + u_4 \end{bmatrix}$$

$$f_1 + f_2 + f_3 + f_4 = \boldsymbol{f}^T \boldsymbol{e} = \boldsymbol{f} \cdot \boldsymbol{e} = \langle \boldsymbol{f}, \boldsymbol{e} \rangle = 0.$$

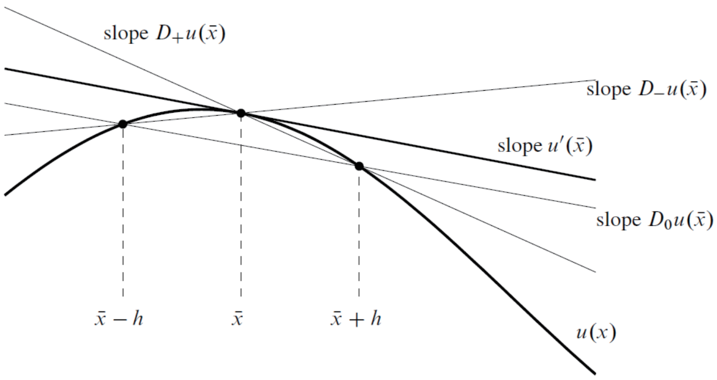

Figure 1: Finite Differences

§1.2 Differences, Derivatives, Boundary Conditions

Observation: -1, 2, -1 produces a second difference.
K_n, C_n, T_n, B_n are all involved in approximating the equation

$$-\frac{d^2u}{dx^2} = f(x)$$

with boundary conditions at $x = 0$ and $x = 1$.

Part I: Finite Differences

-want to approximate $\frac{du}{dx}$.
$\rightsquigarrow \frac{du}{dx} \approx \frac{\triangle u}{\triangle x}$ if $\triangle\ x$ is small.

ex) Choose test function $u(x) = x^2$.

Forward difference $\triangle_+ f(x) = \frac{u(x+h)-u(x)}{h}$
ex) $\frac{(x+h)^2-x^2}{h} = 2x + h$
Backward difference $\triangle_- f(x) = \frac{u(x)-u(x-h)}{h}$
ex) $\frac{x^2-(x-h)^2}{h} = 2x - h$
Centered difference $\triangle_0 f(x) = \frac{u(x+h)-u(x-h)}{2h}$
ex) $\frac{(x+h)^2-(x-h)^2}{h} = 2x$

Taylor series: series in h

$$u(x+h) = u(x) + hu'(x) + h^2/2u''(x) + h^3/3!u'''(x) + \cdots$$
$$u(x-h) = u(x) - hu'(x) + h^2/2u''(x) - h^3/3!u'''(x) + \cdots$$

$$\frac{u(x+h) - u(x)}{h} = u'(x) + h/2u''(x) + \cdots$$

: It is **first order** accurate.

$$\frac{u(x+h) - u(x-h)}{2h} = u'(x) + h^2/3!u'''(x) + \cdots$$

: Centered is **second** order.

Centered difference matrix:

$$\triangle_0 = \begin{bmatrix} \ddots & & & \\ -1 & 0 & 1 & \\ & -1 & 0 & 1 \\ & & & \ddots \end{bmatrix} \begin{bmatrix} u_{i-1} \\ u_i \\ u_{i+1} \\ u_{i+2} \end{bmatrix} = \begin{bmatrix} \vdots \\ u_{i+1} - u_{i-1} \\ u_{i+2} - u_i \\ \vdots \end{bmatrix}$$

Rmk: $\triangle_0^T = -\triangle_0$: antisymmetric (skew symmetric).
The centered difference is the average of forward and backward.

- **Second Differences from First Differences**

$$\triangle^2 u_i = \triangle_- \triangle_+ u_i = 1/h \left[\left(\frac{u_{i+1} - u_i}{h} \right) - \left(\frac{u_i - u_{i-1}}{h} \right) \right] = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

$$\triangle^2 u(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = u''(x) + 2h^2/4!u^{(4)}(x) + \cdots$$

: **second order accuracy**

- **The Important Multiplications**

1. For constant and linear vectors, the second difference are zero:

$$\triangle^2(\text{constant}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

$$\triangle^2(\text{linear}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

For squares, the second difference is constant:

$$\triangle^2(\text{squares}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 9 \\ 16 \end{bmatrix} = \begin{bmatrix} \vdots \\ 2 \\ 2 \\ \vdots \end{bmatrix}$$

2. Second difference of the ramp vector produce the delta vector:

$$\triangle^2(\text{ramp}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \text{delta}$$

3. Second difference of the sine and cosine and exponential produce $2\cos t -$ 2 times those vectors.

$$\triangle^2(\text{sines}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \sin t \\ \sin 2t \\ \sin 3t \\ \sin 4t \end{bmatrix} = (2cost - 2) \begin{bmatrix} \sin t \\ \sin 2t \\ \sin 3t \\ \sin 4t \end{bmatrix}$$

$$\triangle^2(\text{cosines}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \cos t \\ \cos 2t \\ \cos 3t \\ \cos 4t \end{bmatrix} = (2\cos t - 2) \begin{bmatrix} \cos t \\ \cos 2t \\ \cos 3t \\ \cos 4t \end{bmatrix}$$

$$\triangle^2(\text{exponentials}) \begin{bmatrix} \ddots & & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & & \ddots \end{bmatrix} \begin{bmatrix} e^{it} \\ e^{2it} \\ e^{3it} \\ e^{4it} \end{bmatrix} = (2\cos t - 2) \begin{bmatrix} e^{it} \\ e^{2it} \\ e^{3it} \\ e^{4it} \end{bmatrix}$$

Remark: sines or cosines or exponentials are eigenvectors of K, T, B, C with the right boundary conditions.

Part II: Finite Difference Equations

$-\frac{d^2u}{dx^2} = f$ with boundary conditions $u(0) = 0$ and $u(1) = 0$.
Divide the interval $[0, 1]$ into equal pieces of length $h = \triangle x$.

$$\text{unknown } u = \begin{bmatrix} u(h) \\ u(2h) \\ \vdots \\ u(nh) \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ where } h = \frac{1}{n+1}.$$

Finite difference equation: $\boxed{-\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f_i}$

The first and last $(i = 1, i = n)$ involves u_0, u_{n+1}.

Ex) Solve $-\frac{d^2u}{dx^2} = 1$ with $u(0) = 0$ and $u(1) = 0$,
$-\frac{u_{i+1}-2u_i+u_{i-1}}{h^2} = 1$, $u_0 = 0$ and $u_{n+1} = 0$.

Sol) **Complete solution:** $u_{\text{complete}} = u_{\text{particular}} + u_{\text{nullspace}}$

Particular solution: $-\frac{d^2u}{dx^2} = 1$ is solved by $u_{\text{particular}} = -\frac{1}{2}x^2$.

Nullspace solution: $-\frac{d^2u}{dx^2} = 0$ is solved by $u_{\text{nullspace}} = Cx + D$.

$u(x) = -\frac{1}{2}x^2 + Cx + D$.
$u(0) = 0$: $D = 0$.

$u(1) = 0$: $-\frac{1}{2} + C = 0 \Rightarrow C = \frac{1}{2}$.
$u(x) = +\frac{1}{2}x - \frac{1}{2}x^2$.
This is special: the differential and difference equation have the **same solutions**!
$u(x) = -\frac{1}{2}x^2 + \frac{1}{2}x$ and $u_i = \frac{1}{2}(ih - i^2h^2)$.
With $h = 1/4$,

$$Ku = f \text{ leads } 16 \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3/32 \\ 4/32 \\ 3/32 \end{bmatrix}$$

- **A Different Boundary Condition**

Ex) Solve $-\frac{d^2u}{dx^2} = 1$ with $\frac{du}{dx}(0) = 0$ (free end) and $u(1) = 0$.
$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = 1$, $\frac{u_1 - u_0}{h} = 0$ and $u_{n+1} = 0$.

Sol) $u(x) = \frac{1}{2}(1 - x^2)$.
We expect a $O(h)$ error because of the forward difference $\frac{u_1 - u_0}{h}$.
For $n = 3$, $h = 1/4$,

$$1/h^2 \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ gives } h^2 \begin{bmatrix} 6 \\ 5 \\ 3 \end{bmatrix}$$

To have $O(h^2)$ accuracy, see the worked examples 1.2.A in the book.

99% of the difficulties with DE's occurs at the boundary.

§1.3 elimination leads to $K = LDL^T$

Two theme of the book:
1. How to understand equations.
2. how to solve them. - This section's topic.

$Ku = f$:
$u = K^{-1}f$ theorectically, not computationally:
u = inv(K) * f in MATLAB.

Solved by Gaussian elimination - LU decomposition.
If symmetric, $K = LDL^T$ (related to Cholesky factorization).
u = K\f in MATLAB. For multiple f's, [L, U] = lu(K).

Ex) $Ku = f$
$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

$$\rightsquigarrow \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 + 1/2 f_1 \\ f_3 \end{bmatrix}$$

by 2nd row $+1/2\times$ 1st row,

$$\rightsquigarrow \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & 0 & 4/3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 + 1/2 f_1 \\ f_3 + 2/3 f_2 + 1/3 f_1 \end{bmatrix}$$

by 3rd row $+\frac{1}{3/2}$ 2st row.

$$\begin{aligned} 2u_1 - u_2 \qquad\qquad &= f_1, \\ 3/2 u_2 \quad - u_3 &= f_2 + 1/2 f_1, \\ 4/3 u_3 &= f_3 + 2/3 f_2 + 1/3 f_1. \end{aligned}$$

Solution by backsubstitution.

Matrix-vector multiplication $\boldsymbol{Ku}$ as a combination of the columns of $\boldsymbol{K}$:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = u_1 \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}$$

Solving a system $\boldsymbol{Ku} = \boldsymbol{f}$ is exactly the same as finding a combination of the column of $\boldsymbol{K}$ that produces the vector $\boldsymbol{f}$.

Multiplier $l_{ij} = \dfrac{\text{entry to eliminate}}{\text{pivot}} \dfrac{(\text{in row } i)}{(\text{in row } j)}$

The convention: Subtract l_{ij} times the pivot row j from row i. Then the i, j entry is 0.

• Elimination Produces $K = LU$

$$K = LU \quad \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -3/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & 0 & 4/3 \end{bmatrix}$$

$\boldsymbol{L}$ reverses the elimination steps.

Suppose the forward elimination uses the multipliers in $\boldsymbol{L}$ to change the rows of $\boldsymbol{K}$ in to the rows of $\boldsymbol{U}$ (Upper triangular). Then $\boldsymbol{K}$ is factored into $\boldsymbol{L}$ times $\boldsymbol{U}$.

$$\begin{aligned} \boldsymbol{Ku} = \boldsymbol{f}: \quad & \boldsymbol{LUu} = \boldsymbol{f} \Rightarrow \boldsymbol{u} = \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}\boldsymbol{f} \\ & \boldsymbol{Lc} = \boldsymbol{f} \Rightarrow \boldsymbol{c} = \boldsymbol{L}^{-1}\boldsymbol{f} \text{ (forward substitution)} \\ & \boldsymbol{Uu} = \boldsymbol{c} \Rightarrow \boldsymbol{u} = \boldsymbol{U}^{-1}\boldsymbol{c} \text{ (back substitution)} \end{aligned}$$

• Singular Systems

Ex) Circulant $C = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & -1 & -1 \\ 0 & 3/2 & -3/2 \\ 0 & -3/2 & 3/2 \end{bmatrix} \rightsquigarrow$

$\begin{bmatrix} 2 & -1 & -1 \\ 0 & 3/2 & -3/2 \\ 0 & 0 & 0 \end{bmatrix} = U$: The rows are linearly dependent

An invertible matrix has a full set of pivots.

No row exchange to get n pivots: A is invertible and $A = LU$.
Row exchange by P to get n pivots: A is invertible and $PA = LU$.
No way to find n pivots: A is singular, there is no inverse matrix A^{-1}.

Pivoting matrix P: i row $\leftrightsquigarrow$ j row
Composition of $I_{n-2\times n-2}$ except i, j rows and columns and $\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$ for i, j rows and columns.

Ex) 2 row $\leftrightsquigarrow$ 3 row: $\begin{bmatrix} 1 & & & \\ & & 1 & \\ & 1 & & \\ & & & 1 \end{bmatrix}$

- **Symmetry Converts $K = LU$ to $K = LDL^T$**

$\underbrace{K}_{\text{symmetric}} = \underbrace{L}_{\text{lower triangular}} \underbrace{U}_{\text{upper triangular}}$: not symmetric

$$K = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -3/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & 0 & 4/3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -3/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 4/3 \end{bmatrix} \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1 & -3/2 \\ 0 & 0 & 1 \end{bmatrix}$$

$= LDL^T$: symmetric factorization

Remark *1. $A^T CA$ is symmetric if C is symmetric.*

2. *DA: rowwise multiplication by the diagonal entry in D.*

3. *AD: columnwise multiplication by the diagonal entry in D.*

4. *For any rectangular matrix A, the product $A^T A$ is square and symmetric.*

5. *$LDL^T = L\sqrt{D}\sqrt{D}L^T = (L\sqrt{D})(L\sqrt{D})^T := \tilde{L}\tilde{L}^T$ is called Cholesky factorization.*

• The Determinant K_n

$$\begin{aligned} \det K &= \text{by definition of determinant} \sim n! : \text{computationally useless!} \\ &= \det LU = \underbrace{\det L}_{=1} \det U \\ &= \text{product of diagonal entries of } U \\ &= 2 \cdot 3/2 \cdot 4/3 \cdots (n+1)/n = n+1 \end{aligned}$$

The LU decomposition is also a quick way to compute determinant.

Remark *1. If a matrix is tridiagonal, then L and U are bidiagonal.*

2. *If a row/column of K starts with p/q zeros (no elimination needed there), then that low of L/ column of U also starts with p/q zeros.*

3. *Zeros inside the band can unfortunately be "filled in" by elimination - It leads to fundamental problem of reordering the rows and columns to make the p's and q's are as large as possible.*

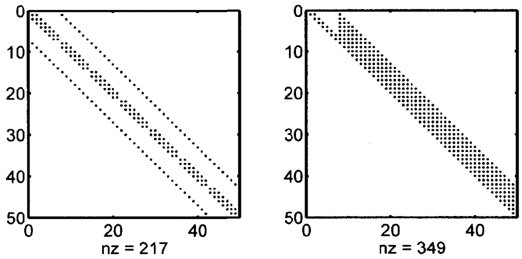

Figure 2: Left: The sparsity pattern of $\boldsymbol{K}$ for 2 dim. Right: The sparsity pattern of the Cholesky factor of $\boldsymbol{K}$.

• Positive Pivots and Positive Determinant

If is positive definite ($x^T A x > 0$ if $x \neq 0$) if all pivots are positive.

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 1 & \\ b/a & 1 \end{bmatrix} \begin{bmatrix} a & \\ & (ac-b^2)/a \end{bmatrix} \begin{bmatrix} 1 & b/a \\ & 1 \end{bmatrix}$$

$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive definite iff $a > 0$ and $ac - b^2 > 0$.

• Operation Counts

LU decomposition $\sim 2/3n^3$ in general, $1/3n^3$ if symmetric.

Operation Count	Full	Banded	Tridiagonal
Factor: Find L and U	$\approx 2/3n^3$	$2w^2n + wn$	$3n$
Solve: Forward and back on f	$2n^2$	$4wn + n$	$5n$

§1.4 Inverse and Delta Functions

: Want to solve for $\boldsymbol{f}$ = point load.

$$\boldsymbol{Ku} = \boldsymbol{\delta}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \text{ (j th entry)} = j \text{ th column of } \boldsymbol{I}.$$

$-u'' = \delta(x-a)$: Green's function u.

Delta function $\delta(x)$:

$\delta(x) = 0$ if $x \neq 0$, $\int_{-\infty}^{\infty} \delta(x)dx = 1$.

: not "true function". "spike", "point load", "impulsive"concentrated at $x = 0$, "infinitely tall and infinitely thin"

A sequence of functions generating or approximating Dirac delta:

$$f_k(x) = \begin{cases} \dfrac{1}{2k}, & \text{if } -k \le x \le k, \\ 0, & \text{otherwise.} \end{cases}$$

$\int_{-\infty}^{\infty} f_k(x)\, dx = 1$ and '$f_k \xrightarrow[k \to 0]{} \delta$ '.

Note that

$$\boldsymbol{K}[\boldsymbol{u}_1|\boldsymbol{u}_2|\cdots|\boldsymbol{u}_n] = [\boldsymbol{\delta}_1|\boldsymbol{\delta}_2|\cdots|\boldsymbol{\delta}_n] \Longleftrightarrow \boldsymbol{K}\boldsymbol{K}^{-1} = \boldsymbol{I}.$$

u_j = column j of K^{-1}. We are solving $KK^{-1} = I$ column by column. If we know the Green's function for all point load $\delta(x-a)$, we can solve $-u'' = f$ for any load $f(x)$.

So, it is the "discrete Green's function".

K_{ij}^{-1} = the solution at point i from a load at point j.

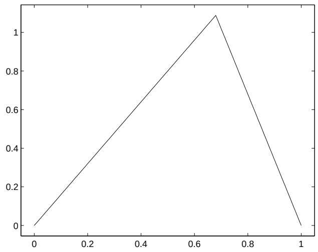

Figure 3: Green's function for fixed-fixed case

• Concentrated load

Ex) $-\frac{d^2u}{dx^2} = \delta(x-a)$ with $\begin{cases} \text{fixed} & u(0)=0 \text{ and fixed } u(1)=0 \\ \text{free} & u'(0)=0 \text{ and fixed } u(1)=0. \end{cases}$

Sol) $\int_{\text{left}}^{\text{right}} -\frac{d^2u}{dx^2}\,dx = \int_{\text{left}}^{\text{right}} \delta(x-a)\,dx$

$\Longrightarrow -\left(\frac{du}{dx}\right)_{\text{right}} + \left(\frac{du}{dx}\right)_{\text{left}} = 1$: The slope drops by 1.

Since $u''=0$ except $x=a$,

$$u = \begin{cases} Ax+B & \text{if } x<a, \\ Cx+D & \text{if } x>a. \end{cases}$$

Boundary Conditions	Jump/No Jump Conditions at $x=a$
fixed $u(0)=0$: $B=0$	No jump in u : $Aa+B=Ca+D$
fixed $u(1)=0$: $C+D=0$	Drop by 1 in u' : $A=C+1$

$$\Longrightarrow \underbrace{u(x)}_{u(x;a)} = \begin{cases} (1-a)x & \text{if } x<a, \\ a(1-x) & \text{if } x>a. \end{cases}$$

Remark $u(x;a)$ *is symmetric w.r.t.* x *and* a. *Note that its discrete version* K^{-1} *is also symmetric (since* K *is symmetric).*

• Delta Function and Green' function

Delta ft $\delta(x)$ $\xleftarrow{\delta(x)=\frac{dS}{dx}}$ / $\xrightarrow{\int_{-\infty}^{x}\delta(y)dy}$ Step ft S(x) $\xleftarrow{S(x)=\frac{dR}{dx}}$ / $\xrightarrow{\int_{-\infty}^{x}S(y)dy}$ Ramp ft $R(x)$

where $S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$ and $R(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$.

The first derivative of the ramp function $R(x)$ jumps by 1 at 0 and the second derivative is a delta function.

Complete solution $-\dfrac{d^2u}{dx^2} = \delta(x-a)$ is solved by

$$u(x) = \underbrace{-R(x-a)}_{\text{particular solution}} + \underbrace{Cx + D}_{\text{null space } u''=0}$$

$0 = u(0) = -R(0-a) + C \cdot 0 + D \Longrightarrow D = 0.$
$0 = u(1) = -R(1-a) + C + D = a - 1 + C \Longrightarrow C = 1 - a.$

$$\underbrace{u(x)}_{u(x,a)} = -R(x-a) + (1-a)x = \begin{cases} (1-a)x & \text{if } x \leq a, \\ (1-x)a & \text{if } x \geq a. \end{cases}$$

The response at x to a load at a equals the response at a to a load at x: symmetric.
cf) $(K^{-1})_{ij} = (K^{-1})_{ji}$
Free-Fixed: $u'(0) = 0, u(1) = 1$

$$\underbrace{u(x)}_{u(x,a)} = -R(x-a) + (1-a)x = \begin{cases} 1-a & \text{if } x \leq a, \\ 1-x & \text{if } x \geq a. \end{cases}$$

• Discrete Vectors: Load and Step and Lamp

The delta vector δ: $\delta = (\cdots, 0, 0, 1, 0, 0, \cdots)$
The step vector S: $S = (\cdots, 0, 0, 1, 1, 1, \cdots)$
The Lamp vector R: $R = (\cdots, 0, 0, 0, 1, 2, \cdots)$
Note that $\boxed{\triangle_- S = \delta \text{ but } \triangle_+ R = S}$
$\triangle^2 = \triangle_- \triangle_+$ so,

$$\boxed{\triangle^2 R = \triangle_- \triangle_+ R = \triangle_- S = \delta}$$

$$\triangle^2(\text{ramp}) \begin{bmatrix} \ddots & & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \text{delta}$$

The solution to $\triangle^2 u = 0$ are "linear vectors" with $u_i = Ci + D$.
The complete solution to $\triangle^2 u = \delta$ is

$$u_i = \underbrace{Ri}_{u_{\text{particular}}} + \underbrace{Ci + D}_{u_{\text{nullspace}}}$$

cf) $u(x) = Rx + Cx + D$
Sampling the ramp $u(x)$ at equally space points without any error.

• The Discrete Equations $Ku = \delta_j$ and $Tu = \delta_j$

$-\triangle^2 u = \delta_j$: $u_i = -R_{i-j} + Ci + D$
$u_0 = -R_{0-j} + C \cdot 0 + D = 0 \Longrightarrow D = 0$
$u_{n+1} = -R_{n+1-j} + C(n+1) + 0 = 0$
$\Longrightarrow C = \frac{n+1-j}{n+1} = \underbrace{1 - \frac{j}{n+1}}_{1-a}$

$$\text{Fixed ends: } u_i = -R_{i-j} + Ci = \begin{cases} \left(\dfrac{n+1-j}{n+1}\right) i & \text{if } i \le j, \\ \left(\dfrac{n+1-i}{n+1}\right) j & \text{if } i \ge j. \end{cases}$$

$$\text{Note: } K_n^{-1} \text{ is symmetric. cf) } u(x) = \begin{cases} 1-a & \text{if } x \le a, \\ 1-x & \text{if } x \ge a. \end{cases}$$

$$\text{Free-Fixed: } u_i = -R_{i-j} + (n+1-j) = \begin{cases} n+1-j & \text{if } i \le j, \\ n+1-i & \text{if } i \ge j. \end{cases}$$

$$\text{cf) } u(x) = \begin{cases} 1-a & \text{if } x \le a, \\ 1-x & \text{if } x \ge a. \end{cases}$$

Green's Function and Inverse Matrix

$$f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = f_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + f_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + f_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

:A combination of n point loads

$$\Rightarrow \boxed{\begin{aligned} K^{-1}f = f_1 \underbrace{(\text{column } 1 \text{ of } K^{-1})}_{K^{-1}\delta_1} &+ f_2 \underbrace{(\text{column } 2 \text{ of } K^{-1})}_{K^{-1}\delta_2} \\ &+ f_3 \underbrace{(\text{column } 3 \text{ of } K^{-1})}_{K^{-1}\delta_3} \end{aligned}}$$

The load $f(x)$ is an integral of point load $f(a)\delta(x-a)$.

$$\boxed{-u'' = f(x) = \int_0^1 f(a)\delta(x-a)da \Longrightarrow u(x) = \int_0^1 f(a)u(x,a)da}$$

The Green's function $u(x,a)$ corresponds to "row x and column a of continuous K^{-1}.

§1.5 Eigenvalues Eigenvectors

Part I: $\boldsymbol{Ax} = \boldsymbol{\lambda x}$ and $\boldsymbol{A^k x} = \boldsymbol{\lambda^k x}$ and Diagonalizing $\boldsymbol{A}$

• Matrix as a Linear Transformation

A matrix $\boldsymbol{A}$ is considered as a linear transformation.

$$\begin{aligned} \boldsymbol{A} : \mathbb{R}^n &\to \mathbb{R}^m \\ \boldsymbol{x} &\mapsto \boldsymbol{Ax} \end{aligned}$$

Definition of linear : $\boldsymbol{A}(r\boldsymbol{x} + s\boldsymbol{y}) = r\boldsymbol{Ax} + s\boldsymbol{Ay}$

'Superposition Principle'

Example 1 $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$: *identity*, $\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$: *dilation*, $\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$: *rotation*, $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$: *reflection.*

• We may regard linear transformation as a composition of those dilations, rotations, and reflections, etc.

• Eigenvalues and Eigenvectors

Definition 2 $\boldsymbol{\lambda}$: *eigenvalue and* $\boldsymbol{x}$: *eigenvector if*

$$\boldsymbol{Ax} = \boldsymbol{\lambda x}, \quad \boldsymbol{x} \neq \boldsymbol{0}.$$

Geometrically, along the eigen-direction, there is only scaling or dilation.

More specifically, the special vector $\boldsymbol{x}$ lies along the same line as $\boldsymbol{Ax}$. The eigenvalue $\boldsymbol{\lambda}$ tells whether the vector $\boldsymbol{x}$ is stretched or shrunk or reversed or left unchanged.

eigen: prime in German

Why are we interested in eigenvalues and eigenvectors?

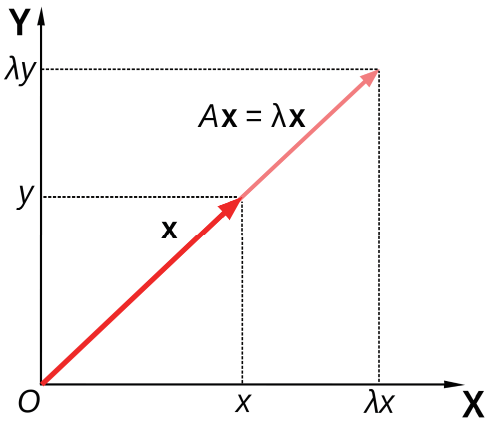

Figure 4: Geometric interpretation of eigenvalue and eigenvector

- It reveals the 'innate or invariant structure' of $\boldsymbol{A}$.
Especially, eigenvalues are invariant under change of basis.
We can understand the matrix $\boldsymbol{A}$ easily by observing the eigenstructure of $\boldsymbol{A}$, called spectrum.

The 'easiest' matrix:

$$\boldsymbol{I} = \begin{bmatrix} \mathbf{1} & & \\ & \ddots & \\ & & \mathbf{1} \end{bmatrix}$$: leave every entry unchanged.

The 'second easiest' matrix:

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{d}_1 & & \\ & \ddots & \\ & & \boldsymbol{d}_n \end{bmatrix}$$

$$\boldsymbol{D} \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{d}_1\boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{d}_n\boldsymbol{x}_n \end{bmatrix}$$: cooridinatewise multiplication.

$\boldsymbol{D}\boldsymbol{x} = \boldsymbol{b}$ can be easily solved by division.
$\boldsymbol{D}$ is positive definite $\Leftrightarrow \boldsymbol{d}_i > \mathbf{0}$ for all $\boldsymbol{i}$.
$\boldsymbol{x}^T\boldsymbol{D}\boldsymbol{x} = \sum_{i=1}^{n} \boldsymbol{d}_i\boldsymbol{x}_i^2 = \boldsymbol{d}_1\boldsymbol{x}_1^2 + \cdots + \boldsymbol{d}_n\boldsymbol{x}_n^2$.

$\boldsymbol{A} \sim \boldsymbol{D}$: A is 'similar' to $\boldsymbol{D}$?
Can we treat or understand $\boldsymbol{A}$ as a diagonal matrix?

Definition 3 *$\boldsymbol{A}$ is similar to $\boldsymbol{B}$ if there exists a nonsingular $\boldsymbol{X}$ such that $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}$.*

Remark *1. If $\boldsymbol{A}$ is similar to a diagonal matrix $\boldsymbol{D}$, then $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{D}$, called diagonalization.*

2. Since $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{D}$, $\boldsymbol{X}$ is the eigenvectors of $\boldsymbol{A}$ with eigenvalues $\boldsymbol{D}$.

3. $\boldsymbol{X}$ is nonsingular means the columns of $\boldsymbol{X}$ consists of a basis.

4. Diagonalization is equivalent to finding $\boldsymbol{n}$ eigenvalues and eigenvectors.

• Diagonalizing a matrix

$\boldsymbol{A}$ is a $\boldsymbol{n}$ by $\boldsymbol{n}$ matrix with $\boldsymbol{n}$ independent eigenvectors $\boldsymbol{x}_1, \cdots \boldsymbol{x}_n$ with eigenvalues $\boldsymbol{\lambda}_1, \cdots \boldsymbol{\lambda}_n$, respectively.

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{X} &= \boldsymbol{A}[\boldsymbol{x}_1|\cdots|\boldsymbol{x}_n] = [\boldsymbol{A}\boldsymbol{x}_1|\cdots|\boldsymbol{A}\boldsymbol{x}_n] = [\boldsymbol{\lambda}_1\boldsymbol{x}_1|\cdots|\boldsymbol{\lambda}_n\boldsymbol{x}_n] \\
&= [\boldsymbol{x}_1|\cdots|\boldsymbol{x}_n] \begin{bmatrix} \boldsymbol{\lambda}_1 & & \\ & \ddots & \\ & & \boldsymbol{\lambda}_n \end{bmatrix} = \boldsymbol{X}\boldsymbol{\Lambda}
\end{aligned}
$$

Now $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \Lambda$. then $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$: Consider $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}\boldsymbol{y}$.
$\boldsymbol{X}^{-1}$: express $\boldsymbol{v}$ w.r.t the eigenbasis $\boldsymbol{X}$.
$\boldsymbol{\Lambda}$: dilation along each eigendirection.
$\boldsymbol{X}$: send back to the original basis.
Thus, the role of $\boldsymbol{X}$ is change of basis.

When square matrices are diagonalizable?

1. In general, it is not guaranteed because there may not be $\boldsymbol{n}$ independent eigenvectors.

   Other decompositions: Jordan canonical form, Schur's canonical form, Singular Value decomposition (SVD).

2. If A is symmetric (for real-valued) or Hermite (for complex-valued), OK.

3. The weakest condition is possibly $AA^T = A^TA$, called normal matrix.

• Symmetric Matrices and Orthonormal Eigenvectors

If A is symmetric, there exist n independent eigenvectors $x_1, \cdots, x_n$ with n real-valued eigenvalues $\lambda_1, \cdots, \lambda_n$. Furthermore, they are orthogonal:

$$\boxed{AU = U\Lambda} \quad \text{and} \quad \boxed{U^TU = I}$$

$$\boxed{A = U\Lambda U^T} \quad \text{or} \quad \boxed{U^TAU = \Lambda}$$

Definition 4 *The vectors $u_i, \cdots u_n$ are orthonormal if*

$$\langle u_i, u_j \rangle = u_i^T u_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \quad \text{(normality)}, \\ 0 & \text{if } i \neq j \quad \text{(orthogonality)}. \end{cases}$$

Definition 5 *The square matrix U is orthogonal if $U^TU = I$. i.e.*

$$U^TU = \begin{bmatrix} u_1^T \\ \hline u_2^T \\ \hline \vdots \\ \hline u_n^T \end{bmatrix} \begin{bmatrix} u_1 & \vline & u_2 & \vline & \cdots & \vline & u_n \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} = I$$

• Finding eigenvalues and eigenvectors

Finding x and λ satisfying $Ax = \lambda x$ is n equations with $n+1$ unknown. We first try to find λ and next find x:

$$\begin{aligned} \lambda : Ax = \lambda x, \quad x \neq 0 &\Leftrightarrow (A - \lambda I)x = 0, \quad x \neq 0 \\ &\Leftrightarrow A - \lambda I \text{ is singluar} \\ &\Leftrightarrow \boxed{\det(A - \lambda I) = 0\text{: charateristic equation}} \end{aligned}$$

$$\begin{aligned} \det(A - \lambda I) &= c_n\lambda^n + c_{n-1}\lambda^{n-1} \cdots + c_0 \\ &= c_n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) \end{aligned}$$

by the Fundamental Theorem of Algebra.

Remark :

1. $\det(A - \lambda I) \sim O(n!)$, not useful computationally.

2. *Furthermore, it is sensitive perturbation and rounding error. Also finding roots is not easy.*

3. *Numerically power method or some other numerical algorithms are used.*

   Ex) `eig(A)` *in MATLAB*

4. *One great success of numerical linear algebra is the development of fast and stable algorithm to compute eigenvalues, especially for the symmetric case.*

Example 6 *Consider the symmetric case* $\boldsymbol{K} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

Sol)

$$\det(\boldsymbol{K} - \lambda \boldsymbol{I}) = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1$$
$$= \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$$

For $\lambda = 1 : \boldsymbol{K} - \boldsymbol{I} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

$(\boldsymbol{K} - \boldsymbol{I})\boldsymbol{x} = (\boldsymbol{K} - \boldsymbol{I})[\boldsymbol{u}; \boldsymbol{v}] = 0 \Rightarrow \boldsymbol{u} - \boldsymbol{v} = 0$

$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *or* $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

For $\lambda = 3 : \boldsymbol{K} - 3\boldsymbol{I} = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$

$(\boldsymbol{K} - 3\boldsymbol{I})\boldsymbol{x} = (\boldsymbol{K} - 3\boldsymbol{I})[\boldsymbol{u}; \boldsymbol{v}] = 0 \Rightarrow \boldsymbol{u} + \boldsymbol{v} = 0$

$\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *or* $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Now

$$\boldsymbol{K} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^T$$

Note that $\boldsymbol{K}$ is a discrete approximation of
$-\frac{d^2u}{dx^2}$ with $\boldsymbol{u}(0) = 0$ and $\boldsymbol{u}(1) = 0$.

$$-\frac{d^2u}{dx^2} = \lambda u \Leftrightarrow \widetilde{K}u = \lambda u \text{ with } \widetilde{K}u = -\frac{d^2u}{dx^2}$$

and

$$\widetilde{K} \text{ is linear}.$$

$$\begin{cases} -\dfrac{d^2u}{dx^2} = \lambda u, \\ u(0) = 0 \text{ and } u(1) = 0. \end{cases}$$

Eigenfunctions: $\sin k\pi x$, $k = 1, 2, \cdots$.
With $h = 1/3$, $x_0 = 0, x_1 = 1/3, x_2 = 2/3, x_3 = 1$,
$\sin \pi x \sim [\sin \pi x_1; \sin \pi x_2] = [\sin 1/3\pi; \sin 2/3\pi] = \frac{1}{\sqrt{2}}[1; 1]$
and
$\sin 2\pi x \sim [\sin 2\pi x_1; \sin 2\pi x_2] = [\sin 2/3\pi; \sin 4/3\pi] = \frac{1}{\sqrt{2}}[1; -1]$.

$$K = U\Lambda U^T, \quad U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

$K^2 = (U\Lambda U^T)(U\Lambda U^T) = U\Lambda^2 U^T$.

$K^n = U\Lambda^n U^T = U\begin{bmatrix} 1 & 0 \\ 0 & 3^n \end{bmatrix}U^T$: K^n grows like 3^n.

The product of n eigenvalues equals the determinant of A.

$$\det A = \prod_{i=1}^{n} \lambda_i$$

Proof) 'determinant of products is product of determinants'.

$$\det A = \det(U\Lambda U^{-1}) = \det U \det \Lambda \det U^{-1}$$
$$= \det(UU^{-1}) \det \Lambda = \det \Lambda.$$

Remark *It is the constant term w.r.t.* λ *in* $\det(A - \lambda I)$.

The sum of the n eigenvalues equal the sum of the n diagonal entries.

$$\text{tr } A = \sum_{i=1}^{n} \lambda_i$$

Remark *It is the coefficient of* $(-\lambda)^{n-1}$ *in* $\det(A - \lambda I)$.

If $A = X\Lambda X^{-1}$ with no nonzero λ_i, then A is invertible and

$$A^{-1} = X\Lambda^{-1}X^{-1} = X \begin{bmatrix} 1/\lambda_1 & & \\ & \ddots & \\ & & 1/\lambda_n \end{bmatrix} X^{-1},$$

which is an eigenvalue decomposition of A^{-1}.

• The Power of a Matrix

-Eigenvalues have their greatest importance in dynamic problems.

Example 7 *Population problem:* $u(t+1) = Au(t)$ *each year where*

$$\begin{bmatrix} u_1(t+1) \\ u_2(t+1) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}.$$

The column sums are always 1, *which means nobody is created or destroyed. Furthermore, populations stay positive because has no negative entries. This type of matrices are called* *Markov matrix* *which expresses probability transition matrix.*

Let $u(0) = \begin{bmatrix} 1000 \\ 0 \end{bmatrix}$. *The matrix* A *has eigenvectors* $\begin{bmatrix} 600 \\ 400 \end{bmatrix}, \begin{bmatrix} 400 \\ -400 \end{bmatrix}$ *with eigenvalues* $1, 1/2$ *respectively.*

If $Ax = \lambda x$, $A^2x = \lambda^2 x, \cdots, A^k x = \lambda^k x$.

Similarly, $A^k(\alpha_1 x_1 + \cdots + \alpha_n x_n) = \alpha_1\lambda_1^k x_1 + \cdots + \alpha_n \lambda_n^k x_n$.

$$\begin{aligned} u(t) = A^t \begin{bmatrix} 600 \\ 400 \end{bmatrix} &= A^t \left(\begin{bmatrix} 600 \\ 400 \end{bmatrix} + \begin{bmatrix} 400 \\ -400 \end{bmatrix} \right) \\ &= \underbrace{1^t \begin{bmatrix} 600 \\ 400 \end{bmatrix}}_{\text{steady state}} + \underbrace{\left(\frac{1}{2}\right)^t \begin{bmatrix} 400 \\ -400 \end{bmatrix}}_{\text{transient}} \xrightarrow{t \to \infty} \begin{bmatrix} 600 \\ 400 \end{bmatrix}. \end{aligned}$$

• Three steps to find $u_k = A^k u_0$ from eigenvalues and eigenvectors

Step 1 : Write $\boldsymbol{u}_0$ as a combination of the eigenvectors:

$$\boldsymbol{u}_0 = \alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_n \boldsymbol{x}_n$$

Step 2 : Multiply each number α_j by $(\lambda_j)^k$.

Step 3 : Recombine the eigenvectors into

$$\boldsymbol{u}_k = \alpha_1 \lambda_1^k \boldsymbol{x}_1 + \cdots + \alpha_n \lambda_n^k \boldsymbol{x}_n$$

In matrix,

Step 1 : $\boldsymbol{u}_0 = [\boldsymbol{x}_1 | \cdots | \boldsymbol{x}_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = X\boldsymbol{a}, \quad \alpha = X^{-1}\boldsymbol{u}_0.$

Step 2 : Multiply $\begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \Lambda^k \alpha, \quad \Lambda^k X^{-1} \boldsymbol{u}_0.$

Step 3 : Recombine $\boldsymbol{u}_k = [\boldsymbol{x}_1 | \cdots | \boldsymbol{x}_n] \begin{bmatrix} \lambda_1^k \alpha_1 \\ \vdots \\ \lambda_n^k \alpha_n \end{bmatrix} = X \Lambda^k \alpha$

$$\boldsymbol{u}_k = X \Lambda^k X^{-1} \boldsymbol{u}_0.$$

Remark

$$A\boldsymbol{y} = X \Lambda X^{-1} \boldsymbol{y}$$

$X^{-1}\boldsymbol{y}$ expresses $\boldsymbol{y}$ w.r.t. the basis $X = [\boldsymbol{x}_1 | \cdots | \boldsymbol{x}_n]$.
Λ multiplies eigenvalues.
X recombines.

- **Application to Vector Differential equations**

Example 8

$$\begin{cases} \dfrac{dy}{dt} = ay & \textit{general solution } y(t) = Ce^{at}. \\ y(0) = y_0 & \textit{It determines } C. \end{cases}$$

The solution $y(t) = y_0 e^{at}$ *decays if* $a < 0$*: stability.*
The solution $y(t) = y_0 e^{at}$ *grows if* $a > 0$*: instability.*
When a *is a complex number, its real part determines the growth or decay, the imaginary part gives oscillatory factor since*

$$e^{iwt} = \cos wt + i \sin wt \quad \textit{(Euler formula)}$$

Vectorial case or system of equations

$$\frac{dy}{du} = Au, \quad u(0) = u_0$$

Example 9 $\begin{aligned} \frac{dy}{dt} &= 2y - z \\ \frac{dz}{dt} &= -y + 2z \end{aligned}$, $\quad \frac{d}{dt}\begin{bmatrix} y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{K_2} \begin{bmatrix} y \\ z \end{bmatrix}$

Sol) Assume $u(t) = e^{\lambda t} x = e^{\lambda t} \begin{bmatrix} y \\ z \end{bmatrix}$.

$\frac{du}{dt} = \lambda e^{\lambda t} x = LHS = RHS = Ku = e^{\lambda t} K x$

$$Kx = \lambda x : \textit{eigenvalue problem}$$

$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *with* $\lambda = 1, 3$ *respectively.*
General solution:

$$\begin{aligned} u(t) =& c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 \\ =& c_1 e^{t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \end{aligned}$$

$$u_0 = u(0) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{X} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

Three steps for powers apply here too:
Expand $u_0 = X\alpha$*:* $\alpha = X^{-1} u_0$.

Multiply each α_j *by* $e^{\lambda_j t}$*:* $\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} a = e^{\Lambda t} X^{-1} u_0$.

Recombine into $u(t) = X e^{\Lambda t} X^{-1} u_0$.

Part II: Eigenvectors for Derivatives and Differences

$$-\frac{d^2u}{dx^2} = \lambda u \text{ is solved by } y = \cos\omega x,\ y = \sin\omega x \text{ with } \lambda = \omega^2.$$

Analog to K_n: fixed-fixed case $y(0) = 0, y(1) = 0$.
$y(x) = \sin k\pi x$ with $\lambda = k^2\pi^2$, $k = 1, 2, \cdots$
Sol) $y = a\cos\omega x + b\sin\omega x$
$0 = y(0) = a$
$0 = y(1) = b\sin\omega \Rightarrow \omega = k\pi$: determined by boundary condition!

Analog to B_n: free-free case $y'(0) = 0, y'(1) = 0$.
$y(x) = \cos k\pi x$ with $\lambda = k^2\pi^2$, $k = 0, 1, 2, \cdots$

Analog to C_n: periodic case $y(0) = y(1), y'(0) = y'(1)$.
$y(x) = \cos 2k\pi x, \sin 2\tilde{k}\pi x$ with $\lambda = 4k^2\pi^2$, $k = 0, 1, 2, \cdots$, $\tilde{k} = 1, 2, \cdots$

Analog to T_n: free-fixed case $y'(0) = 0, y(1) = 0$.
$y(x) = \cos(k + 1/2)\pi x$ with $\lambda = k^2\pi^2$, $k = 0, 1, 2, \cdots$

- **Eigenvectors of K_n: Discrete Sines**

$$-\begin{bmatrix} \sin(j-1)\theta \\ \cos(j-1)\theta \end{bmatrix} + 2\begin{bmatrix} \sin j\theta \\ \cos j\theta \end{bmatrix} - \begin{bmatrix} \sin(j+1)\theta \\ \cos(j+1)\theta \end{bmatrix} = (2 - 2\cos\theta)\begin{bmatrix} \sin j\theta \\ \cos j\theta \end{bmatrix}$$

These are the imaginary and real parts of

$$-e^{i(j-1)\theta} + 2e^{ij\theta} - e^{i(j+1)\theta} = (2 - e^{-i\theta} - e^{i\theta})e^{ij\theta}.$$

The boundary rows decide θ everything!
For $\sin j\theta$,
From the first row,
$2\sin\theta - \sin 2\theta = (2 - 2\cos\theta)\sin\theta$, it is true for any θ.
From the last row,
$-\sin(n-1)\theta - 2\sin n\theta = (2 - \cos\theta)\sin n\theta$,
$-(\sin n\theta\cos\theta - \cos n\theta\sin\theta) + 2\sin n\theta = 2\sin n\theta - 2\sin n\theta\cos\theta$
$\sin n\theta\cos\theta + \cos\theta\sin n\theta = \sin n(\theta + 1) = 0$
$\Rightarrow \theta = \frac{k}{n+1}\pi, k = 1, 2, \cdots$
For $\cos j\theta$,
From the first row,
$2\cos\theta - \cos 2\theta = (2 - 2\cos\theta)\cos\theta$
$\cos^2\theta - 1 = 2\cos^2\theta$
$\Rightarrow -1 = 0$: There is no such θ.

The first eigenvector y_1 will sample the first eigenfunction $y(x) = \sin\pi x$ at n meshpoint with $h = \frac{1}{n+1}$:
First eigenvector = dicrete sine $y_1 = (\sin\pi h, \sin 2\pi h, \cdots, \sin n\pi h)$
First eigenvalue of K_n:
$\lambda_1 = 2 - 2\cos\pi h = 2 - 2(1 - \frac{\pi^2 h^2}{2} + \cdots) \approx \pi^2 h^2$
To match differences with derivatives, divide K by $h^2 = (\triangle x)^2$.

eigenvectors = discrete sines $y_k = (\sin k\pi h, \cdots, \sin nk\pi h)$
eigenvalues of $K_n : 2 - 2\cos k\pi h, \quad k = 1, \cdots n.$

Discrete sine transform

$$\text{DST} = \begin{bmatrix} \sin\frac{\pi}{4} & \sin\frac{2\pi}{4} & \sin\frac{3\pi}{4} \\ \sin\frac{2\pi}{4} & \sin\frac{4\pi}{4} & \sin\frac{6\pi}{4} \\ \sin\frac{3\pi}{4} & \sin\frac{6\pi}{4} & \sin\frac{9\pi}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 1 & 0 & -1 \\ \frac{1}{\sqrt{2}} & -1 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$Q = \frac{1}{\sqrt{2}}$DST is orthogonal, i.e. $Q^T Q = I$, $Q^{-1} = Q^T$.

Remark $\int_0^1 \sin n\pi x \sin m\pi x \, dx = 0$ *if* $n \neq m$.

• Eigenvectors of B_n: Discrete Cosines

eigenvalues of B_n : $2 - 2\cos\frac{k\pi}{n}, \quad k = 0, \cdots n-1$

eigenvectors: $\quad y_k = \left(\cos\frac{1}{2}\frac{k\pi}{n}, \cos\frac{3}{2}, \frac{k\pi}{n} \cdots, \cos\left(n - \frac{1}{2}\right)\frac{k\pi}{n}\right)$

Eigenvalues of B sample $\cos k\pi x$ at the n midpoints $x = (j - \frac{1}{2})/n$.

$$y'(0) = 0 \sim y(x_1) - y(\underbrace{x_0}_{\text{ghost grid}}) = 0$$

$$y'(1) = 0 \sim y(\underbrace{x_{n+1}}_{\text{ghost grid}}) - y(x_n) = 0$$

Since the cosine is even, those vectors have zero slope at the ends:
$\cos -\frac{1}{2}\frac{k\pi}{n} = \cos\frac{1}{2}\frac{k\pi}{n}$ and $\cos\left(n - \frac{1}{2}\right)\frac{k\pi}{n} = \cos\left(n + \frac{1}{2}\right)\frac{k\pi}{n}$
: The reason for choosing midpoints as gridpoints.
Note that $k = 0$ gives the all-ones eigenvector $y_0 = (1, 1, \cdots, 1)$ with $\lambda = 0$: DC vector with zero frequency.

Discrete cosine transform

$$\text{DCT} = \begin{bmatrix} \cos 0 & \cos\frac{1}{2}\frac{\pi}{3} & \cos\frac{1}{2}\frac{2\pi}{3} \\ \cos 0 & \cos\frac{3}{2}\frac{\pi}{3} & \cos\frac{3}{2}\frac{2\pi}{3} \\ \cos 0 & \cos\frac{5}{2}\frac{\pi}{3} & \cos\frac{5}{2}\frac{2\pi}{3} \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2}\sqrt{3} & \frac{1}{2} \\ 1 & 0 & -1 \\ 1 & -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$$

• Eigenvectors of C_n: Powers of $\omega = e^{\frac{2\pi i}{n}}$

Eigenvectors of C_n: Both sine and cosine

$$\text{(Euler formula)} \quad e^{i\theta} = \cos\theta + i\sin\theta$$

$$\text{Circulant matrix (periodic)} \quad C_4 = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

It has constant diagonals with wrap-around.
The k th eigenvector of C_n comes from sampling $y_k(x) = e^{i2\pi kx}$ at the n meshpoints $x = j./n$, $j = 0, \cdots, n-1$.
j th component of y_k: $e^{i2\pi k(j/n)} = \omega^{jk}$ where $\omega = e^{i2\pi/n} = n$ th root of 1.

eigenvalues of $C_n : 2 - \omega^k - \omega^{-k} = 2 - 2\cos\frac{k\pi}{n}, \quad k = 0, \cdots n-1$

eigenvectors: $\quad y_k = (1, \omega^k, \omega^{2k} \cdots, \omega^{(n-1)k})$

• The Fourier Matrix

Discrete Fourier transform (DFT)

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}, \quad (F_n)_{jk} = \omega^{jk} = e^{i2\pi jk/n}$$

The columns are orthogonal in $\mathbb{C}$: $\langle x, y\rangle = x^* y = \bar{x}^T y$.
$\bar{F_4}^T F_4 = 4I$ so that $F_4^{-1} = \frac{1}{4}\bar{F_4}^T$.

In general,

$$\bar{F_n}^T F_n = nI \text{ and } F_n^{-1} = \frac{1}{n}\bar{F_n}^T = \frac{1}{n}F_n^*$$

.

$U_n = \frac{1}{\sqrt{n}}F_n$: The normalized Fourier matrix is unitary.

Columns are 'orthonormal' in $\mathbb{C}$: $\bar{U_n}^T U_n = \frac{1}{\sqrt{n}}\bar{F_n}^T \frac{1}{\sqrt{n}}F_n = I$.

Unitary matrix ($Q^*Q = \bar{Q}^T Q = I$) is the complex analog of orthogonal matrix ($A^T A = I$).

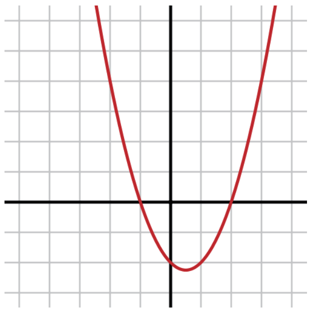

Figure 5: Quadratic function in 1-d example: $x^2 - x - 2$

§1.6 Positive Definite Matrix

What is 'positive definite'?

3 basic facts

1. $K = A^T A$ is symmetric and positive definite (or at least semidefinite).
   $x^T A^T A x = (Ax)^T Ax$
2. If K_1 and K_2 are positive definite, then so is $K_1 + K_2$.
3. All pivots and all eigenvalues of a positive definite matrix is positive.

Why do we want to consider positive definite matrices?
- It is closely related to the concept of energy as the quadratic form $\frac{1}{2}u^T K u$ and we are interested in its minimum.

Example 10 *1 dimensional example (See Fig. 5)*
$f(x) = \frac{1}{2}ax^2 - bx + \underbrace{c}_{=0} = \frac{1}{2}a(x - \frac{b}{a})^2 - \frac{1}{2}\frac{b^2}{a}.$
Optimization: For its minimum,
The first necessary condition: $f'(x) = ax - b$.
The second sufficient condition: $f''(x) = a > 0$.

If K is positive definite,

the minimum of $P(u) = \dfrac{1}{2}u^T K u - u^T f$ is $P_{\text{min}} = -\dfrac{1}{2} f K^{-1} f$ when $Ku = f$.

Example 11 *2 dimensional examples (See Fig. 6)*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, D = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}.$$

$x^T A x = [x\ y] \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 2y^2$: *positive definite (elliptic).*

Similarly,

$x^T B x = x^2$: *semipositive definite (parabolic),*

$x^T C x = x^2 - 2y^2$: *indefinite (hyperbolic),*

$x^T D x = -x^2 - 2y^2 = -(x^2 + 2y^2)$: *negative definite.*

• Examples and Energy-based Definition

Quadratic function:

$$u^T S u = [u_1, u_2] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = a u_1^2 + 2b u_1 u_2 + c u^2.$$

Example 12 *Sum of squares examples (See also Fig. 6)*

Positive definite	*Semipositive definite*	*Indefinite*
$K = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$	$B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$	$M = \begin{bmatrix} 1 & -3 \\ -3 & 1 \end{bmatrix}$
$2u_1^2 - 2u_1 u_2 + 2u_2^2$	$u_1^2 - 2u_1 u_2 + u_2^2$	$2u_1^2 - 6u_1 u_2 + 2u_2^2$
Always positive	*Positive or zero*	*Positive or negative*

$$\begin{aligned} & 2u_1^2 - 2u_1 u_2 + 2u_2^2 \\ =& u_1^2 + (u_1 - u_2)^2 + u_2^2 : A^T A \\ =& 2(u_1 - \frac{1}{2} u_2)^2 + \frac{3}{2} u_2^2 : LDL^T. \end{aligned}$$

$$K = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} = A^T A$$

$$= \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} = LDL^T.$$

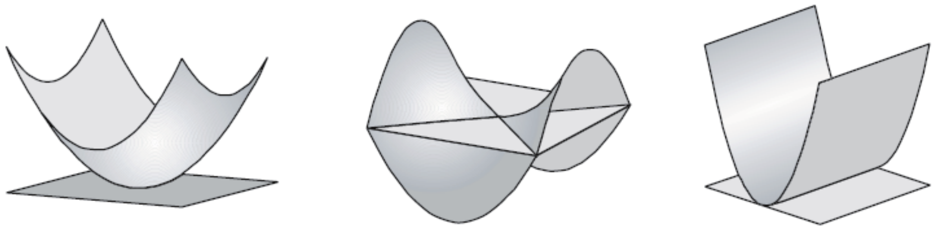

Figure 6: Positive definite, Indefinite, Semidefinite functions in **2** dim

$$u_1^2 - 2u_1u_2 + u_2^2 = (u_1 - u_2)^2.$$
$$2u_1^2 - 6u_1u_2 + 2u_2^2 = (u_1 - 3u_2)^2 - 8u_2^2.$$

- **Positive definiteness from $A^TA, A^TCA, LDL^T, Q\Lambda Q^T$**

1. $K = A^TA$ is symmetric positive definite iff A has independent columns:
$\boxed{u^TKu = u^TA^TAu = (Au)^TAu = \|Au\| > 0}$ for $x \neq 0$ if A has full rank.

2. $K = A^TCA$ is symmetric positive definite iff A has independent columns and C is symmetric positive definite:
$\boxed{u^TKu = u^TA^TCAu = (Au)^TCAu > 0}$ for $x \neq 0$.

3. If symmetric K has a full set of positive pivots, it is positive definite:
$K = LDL^T$, the diagonal pivot matrix D is positive definite and L^T has independent columns.

4. If a symmetric K has all positive eigenvalues in Λ, it is positive definite:
$K = Q\Lambda Q^T$, $Q^{-1} = Q^T$.

- **Minimum Problem in n Dimensions**

Very often, $\frac{1}{2}u^TKu$ is the "internal energy" in the system.

$$P(u) = \frac{1}{2}u^T K u - u^T f : \text{ total energy}$$

$\nabla P = Ku - f = 0$: the first necessary condition

$H = K > 0$, positive definite: the second sufficient conditon

$P(K^{-1}f) = \frac{1}{2}(K^{-1}f)^T K(K^{-1}f) - (K^{-1}f)^T f = -\frac{1}{2}f^T K^{-1} f$

$$\begin{aligned} P(u) - P(K^{-1}f) &= \frac{1}{2}u^T K u - u^T f - (-\frac{1}{2}f^T K^{-1} f) \\ &= \frac{1}{2}(u - K^{-1}f)^T K(u - K^{-1}f) \geq 0. \end{aligned}$$

• Test for a minimum: Positive Definite Second Derivatives

Test for 1 dimensional function:

$$f(x) = \underbrace{f(a)}_{\text{const}} + \underbrace{f'(a)}_{\text{slope}=0}(x-a) + \frac{1}{2}\underbrace{f''(a)}_{\text{concavity}>0}(x-a)^2 + \cdots$$

by Taylor series.

Test for n dimensional function:

$$P(u) = P(u^*) + (u-u^*)^T \nabla P(u^*) + \frac{1}{2}(u-u^*)^T H(u^*)(u-u^*) + \cdots$$

again by Taylor series.

For minimum,

$$\text{(1st derivative vector: slope) } \nabla P(u^*) = \begin{bmatrix} \frac{\partial P}{\partial u_1} \\ \vdots \\ \frac{\partial P}{\partial u_n} \end{bmatrix} = 0$$

and

$$\text{(2nd derivative matrix: concavity) } H_{ij} = \frac{\partial^2 P}{\partial u_i \partial u_j} = \frac{\partial^2 P}{\partial u_j \partial u_i} = H_{ji} > 0$$

i.e. positive definite.

• Newton method

Approximation by by quadratic form:

$$\begin{aligned}
P(u) \approx & \underbrace{P(u^*)}_{=0} + \underbrace{(u-u^*)^T}_{u^T} \underbrace{\nabla P(u^*)}_{-f} + \frac{1}{2} \underbrace{(u-u^*)^T}_{u^T} \underbrace{H(u^*)}_{K} \underbrace{(u-u^*)}_{u} \\
:= & \frac{1}{2} u^T K u - u^T f \\
& \Rightarrow Ku = f.
\end{aligned}$$

$$H(u^*)(u-u^*) = -\nabla P(u^*).$$

Newton's method: $H(u^i)(u^{i+1} - u^i) = -\nabla P(u^i)$

If u^i hits exactly a minimum u^* (not too likely) $\nabla P(u^*) = 0$, so $u^{i+1} - u^i = 0$, no more steps.
It is an iterative method to solve a minimum problem.

§1.7 Numerical Linear Algebra: LU, QR, SVD

Ex) $\boldsymbol{Ku} = \boldsymbol{f}$ or $\boldsymbol{Kx} = \boldsymbol{\lambda x}$ or $\boldsymbol{Mu''} + \boldsymbol{Ku} = \boldsymbol{0}$
Crucial properties of $\boldsymbol{K}$: symmetric? banded? sparse? well-conditioned?

• Three Essential Factorization

1. $\boldsymbol{A} = \boldsymbol{LU}$ = lower triangle matrix × upper triangle matrix
   by Gaussian elimination.
2. $\boldsymbol{A} = \boldsymbol{QR}$ = Orthogonal matrix × upper triangle matrix
   by Gram-Schmidt orthogonalization or Householder transformation.
3. $\boldsymbol{A} = \boldsymbol{U\Sigma V^T}$ = orthonormal columns × singular values × orthonormal lows
   by singular value decomposition.
   It is a 'generalized eigenvalue decomposition'. cf) $\boldsymbol{Q\Lambda Q^T}$.

• Orthogonal Matrices

$\langle \boldsymbol{q_i}, \boldsymbol{q_j} \rangle = \boldsymbol{q_i^T q_j} = \boldsymbol{0}$ if $\boldsymbol{i} \neq \boldsymbol{j}$ (orthogonality)
$\langle \boldsymbol{q_i}, \boldsymbol{q_i} \rangle = \boldsymbol{q_i^T q_i} = \boldsymbol{1}$ (normalization to unit vector)
Let $\boldsymbol{Q} = [\boldsymbol{q_1}|\boldsymbol{q_2}|\cdots|\boldsymbol{q_n}]$.

$$\boldsymbol{Q^T Q} = \begin{bmatrix} \boldsymbol{q_1^T} \\ \hline \boldsymbol{q_2^T} \\ \hline \vdots \\ \hline \boldsymbol{q_n^T} \end{bmatrix} \left[\begin{array}{c|c|c|c} \boldsymbol{q_1} & \boldsymbol{q_2} & \cdots & \boldsymbol{q_n} \end{array} \right] = \begin{bmatrix} \boldsymbol{1} & & \\ & \ddots & \\ & & \boldsymbol{1} \end{bmatrix} = \boldsymbol{I}$$

The inverse is its transpose: $\boldsymbol{Q^{-1}} = \boldsymbol{Q^T}$.
Length preserving (also angle preserving): $\|\boldsymbol{Qx}\| = \|\boldsymbol{x}\|$

Ex) permutations, rotations, reflections.

Example 13 *Permutation: the same rows as $\boldsymbol{I}$, in a different order.*

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix} \qquad \boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{I}$$

Example 14 *Rotation*

Rotation matrix in the $1-3$ *plane:* $\begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$

Example 15 *Reflection: The reflection takes $\boldsymbol{v}$ to $\boldsymbol{Hv}$ on the other side of a plane mirror. The unit vector $\boldsymbol{u}$ perpendicular to the mirror is reversed into $\boldsymbol{Hu} = -\boldsymbol{u}$.*

Reflection matrix $\boldsymbol{u} = (\cos\theta, 0, \sin\theta)$:

$$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} \cos\theta \\ 0 \\ \sin\theta \end{bmatrix} [\cos\theta, 0, \sin\theta]$$

$$= \begin{bmatrix} 1-\cos^2\theta & 0 & -2\sin\theta\cos\theta \\ 0 & 1 & 0 \\ -2\sin\theta\cos\theta & 0 & 1-\sin^2\theta \end{bmatrix} = \begin{bmatrix} -\cos 2\theta & 0 & -\sin 2\theta \\ 0 & 1 & 0 \\ -\sin 2\theta & 0 & \cos 2\theta \end{bmatrix}$$

$\det \boldsymbol{H} = -1$, $\boldsymbol{Hu} = (\boldsymbol{I} - 2\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{u} = \boldsymbol{u} - 2\boldsymbol{u} = -\boldsymbol{u}$.

It is a popular method for QR decomposition.

- **Orthogonalization $\boldsymbol{A} = \boldsymbol{QR}$**

1. Gram-Schmidt algorithm

$$\boldsymbol{A}_{m\times n} = \begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix}$$

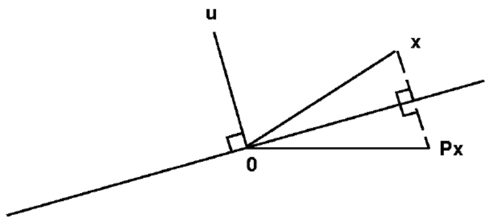

Figure 7: Householder transformation

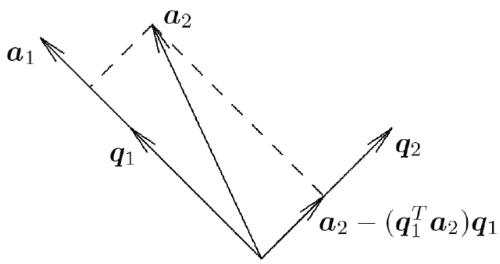

Figure 8: Gram-Schmidt orthogonalization

rank n, independent n vectors, consists of a basis for the column space of A
$q_1 := \frac{a_1}{\|a_1\|}$, $a_1 = r_{11} q_1$ with $r_{11} =: \|a_1\|$.
$B = a_2 - (q_1^T a_2) q_1$ is orthogonal to q_1.
$q_1^T a_2$: projection in the q_1 direction $:= r_{12}$.
$q_2 = \frac{B}{\|B\|} = r_{22}$.

Gram-Schmidt

$$\left[\begin{array}{c|c} a_1 & a_n \end{array}\right] = \left[\begin{array}{cc} q_1 & q_2 \end{array}\right] \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}$$ i.e. $$\begin{aligned} a_1 &= r_{11} q_1 \\ a_2 &= r_{12} q_1 + r_{22} q_2 \end{aligned}$$

2. Householder algorithm : $I - 2uu^T$
- used in MATLAB and popular numerical linear algebra package.

The great virtue of Q is its stability.
$Qx = b$ is perfectly conditioned since $\|x\| = \|b\|$ and an error $\triangle b$ produces an error $\triangle x$ of the same size:
$Q(x + \triangle x) = b + \triangle b$ gives $Q(\triangle x) = \triangle b$ and $\|\triangle x\| = \|\triangle b\|$.

Singular Value Decomposition

Motivation: If $A_{m\times n}$ is symmetric positive definite, $Q\Lambda Q^T$.
If a full-rank matrix A is not symmetric, furthermore, not square i.e. general $m \times n$, what can we do?

$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m}\underbrace{\Sigma}_{m\times n}\underbrace{V^T}_{n\times n} \text{ with } U^TU = I \text{ and } V^TV = I.$$

Note that A^TA is symmetric positive definite.

$$\underbrace{A^TA}_{n\times n} = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^TU^TU\Sigma V^T$$

$$= \underbrace{V}_{n\times n}\underbrace{\Sigma^T\Sigma}_{n\times n}\underbrace{V^T}_{n\times n} = V\begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2\end{bmatrix}V^T$$

$$:= Q\Lambda Q^T = Q\begin{bmatrix}\lambda & & \\ & \ddots & \\ & & \lambda\end{bmatrix}Q^T \text{ with } Q = V, \lambda_i = \sigma^2.$$

It is an eigenvalue decomposition of A^TA.
From $AV = U\Sigma$ $Av_i = \sigma_i u_i \Rightarrow u_i = Av_i/\sigma_i$.
Since $AA^T = (U\Sigma V^T)(U\Sigma V^T)T = U\Sigma\Sigma^TU^T$,
u_i are orthonormal eigenvectors of AA^T.

Reduced SVD :
(rank r case)

$$A_{m\times n} = \underbrace{U_{m\times r}}_{\text{left singular vector}} \Sigma_{r\times r} \underbrace{U^T_{r\times n}}_{\text{right singular vector}}$$

$$= \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix}\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}\begin{bmatrix} v_1^T \\ \vdots \\ v_r^T\end{bmatrix}$$

with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$.
MATLAB command: `svd(A, 0)`.
To complete v's, add any orthogonal basis $v_{r+1}, \cdots, v_n$ for nullspace of A.

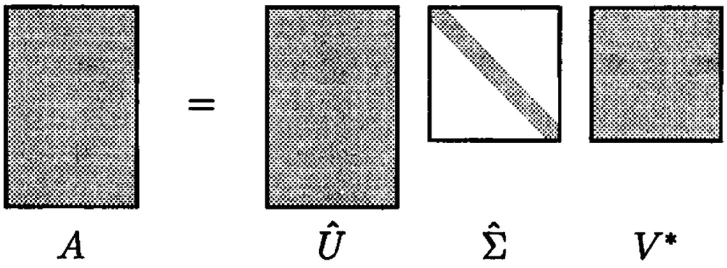

Figure 9: Reduced SVD of the full-rank $\boldsymbol{A}$

To complete $\boldsymbol{u}$'s, add any orthogonal basis $\boldsymbol{u}_{r+1}, \cdots, \boldsymbol{u}_m$ for nullspace of $\boldsymbol{A}^T$.
To complete $\boldsymbol{\Sigma}$ to an $\boldsymbol{m}$ by $\boldsymbol{n}$ matrix, add zeros.

Full SVD :

$$\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\boldsymbol{U}^T_{n\times n}$$

$$= \left[\begin{array}{c|c|c|c|c|c} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r & \boldsymbol{u}_{r+1} & \cdots & \boldsymbol{u}_m \end{array} \right] \begin{bmatrix} \boldsymbol{\sigma}_1 & & \\ & \ddots & \\ & & \boldsymbol{\sigma}_r \\ & & \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1^T \\ \hline \vdots \\ \hline \boldsymbol{v}_r^T \\ \hline \boldsymbol{v}_{r+1}^T \\ \hline \vdots \\ \hline \boldsymbol{v}_n^T \end{bmatrix}$$

MATLAB command: `svd(A)`.

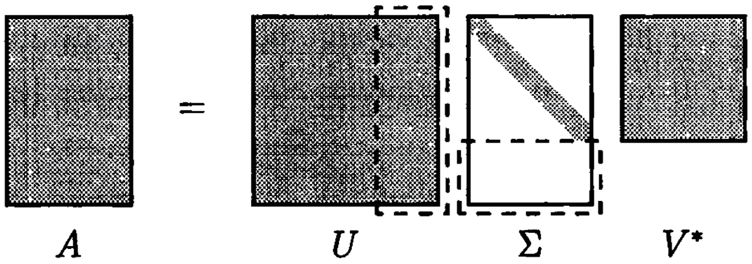

Figure 10: Full SVD of the full-rank $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{u}_1\boldsymbol{\sigma}_1\boldsymbol{v}_1 + \boldsymbol{u}_2\boldsymbol{\sigma}_2\boldsymbol{v}_2 + \cdots \boldsymbol{u}_r\boldsymbol{\sigma}_r\boldsymbol{v}_r$$

$$\boldsymbol{A}\boldsymbol{v}_j = \begin{cases} \boldsymbol{\sigma}_j\boldsymbol{u}_j & \text{for } j \le r, \\ 0 & \text{for } j > r. \end{cases} \qquad \boldsymbol{A}^T\boldsymbol{u}_j = \begin{cases} \boldsymbol{\sigma}_j\boldsymbol{v}_j & \text{for } j \le r, \\ 0 & \text{for } j > r. \end{cases}$$

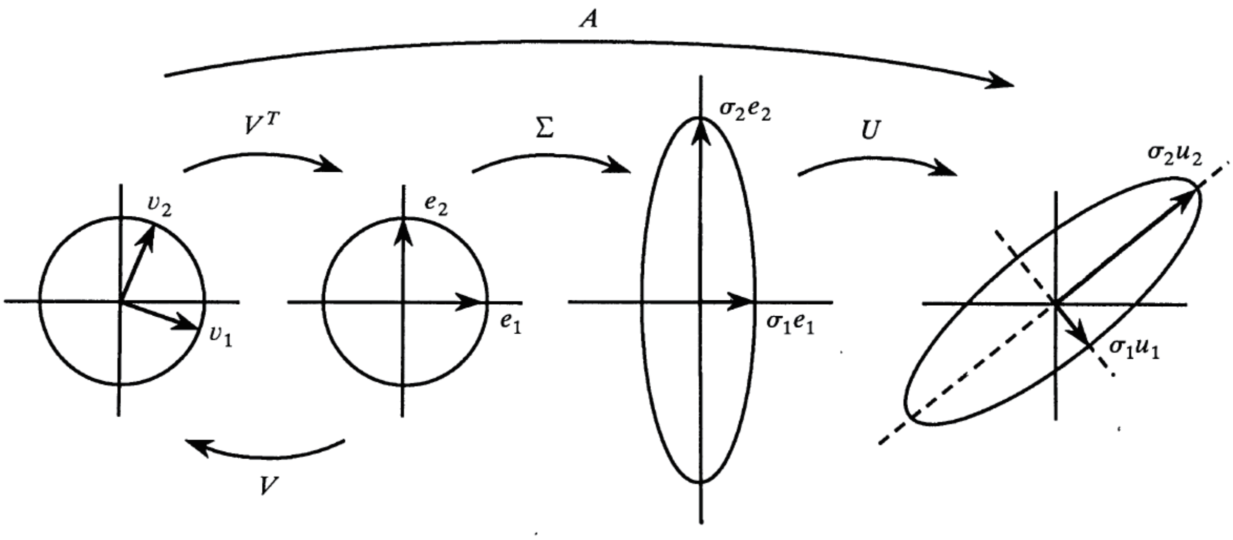

Figure 11: The transformation $\boldsymbol{A}$ in terms of SVD: 'Fundamental Theorem of Linear Algebra' by Strang.

Example 16 *Find the SVD for* $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 7 & 7 \end{bmatrix}$.

Sol) $\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} 1 & 7 \\ 1 & 7 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 7 & 7 \end{bmatrix} = \begin{bmatrix} 50 & 50 \\ 50 & 50 \end{bmatrix}$.

$$\det(\boldsymbol{A}^T\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} 50-\lambda & 50 \\ 50 & 50-\lambda \end{vmatrix} = (50-\lambda)^2 - 50^2$$
$$= \lambda^2 - 100\lambda = \lambda(\lambda - 100).$$

$\lambda = 100, 0$ *with eigenvectors* $[\boldsymbol{v}_1, \boldsymbol{v}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Now $\boldsymbol{\sigma} = 10, 0$:

$$\boldsymbol{u}_1 = \boldsymbol{A}\boldsymbol{v}_1/\boldsymbol{\sigma}_1 = \frac{1}{10}\begin{bmatrix} 1 & 7 \\ 1 & 7 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{10\sqrt{2}}\begin{bmatrix} 2 \\ 14 \end{bmatrix} = \frac{1}{5\sqrt{2}}\begin{bmatrix} 1 \\ 7 \end{bmatrix}.$$

$\boldsymbol{u}_2 = \frac{1}{5\sqrt{2}}\begin{bmatrix} -7 \\ 1 \end{bmatrix}$ *since* $\boldsymbol{u}_2^T\boldsymbol{u}_1 = 0$ *and* $\|\boldsymbol{u}_2\| = 1$.

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \frac{1}{5\sqrt{2}}\begin{bmatrix} 1 & -7 \\ 7 & 1 \end{bmatrix}\begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \textit{(full)}$$
$$= \frac{1}{5\sqrt{2}}\begin{bmatrix} 1 \\ 7 \end{bmatrix} 10 \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix} \quad \textit{(reduced)}$$

Example 17 *SVD of the $n+1$ by n backward difference matrix $\triangle_-$. v_k and u_k are DST and DCT matrices. i.e. $\triangle_- = (DCT)\Sigma(DST)^T$. Thus $\triangle_-(DST)^T = (DCT)\Sigma$.*
cf)$(\sin n\pi x)' = n\pi \cos n\pi x$.

• The Pseudoinverse

If $A = Q\Lambda Q^T = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} Q^T$ with full lank, $AQ = Q\Lambda$.

$A^{-1} = Q\Lambda^{-1}Q^T = Q \begin{bmatrix} 1/\lambda_1 & & \\ & \ddots & \\ & & 1/\lambda_n \end{bmatrix} Q^T$, $A^{-1}Q = Q\Lambda^{-1}$.

If $Aq_i = \lambda q_i$, $A^{-1}q_i = 1/\lambda_i q_i$.
Similarly, $Av_i = \sigma_i u_i$, then $A^{-1}u_i = 1/\sigma_i v_i$.
For a square and invertible A,
if $A = U\Sigma V^T$ then $A^{-1} = V\Sigma^{-1}U^T$.
Now, if A is nonsquare or singular?

Pseudoinverse $\quad A^+ = V\Sigma^+U^T, \quad A^+u_i = \begin{cases} \dfrac{v_i}{\sigma_i} \text{ for } & i \le r \\ 0 \text{ for } & i > r \end{cases}$

Example 18 *Find the pseudoinverse A^+ of $A = \begin{bmatrix} 1 & 1 \\ 7 & 7 \end{bmatrix}$.*

Sol)

$$A^+ = V\Sigma^+U^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1/10 & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{5\sqrt{2}}\begin{bmatrix} 1 & 7 \\ -7 & 1 \end{bmatrix} \quad \textit{(full)}$$

$$= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} 1/10 \frac{1}{5\sqrt{2}}\begin{bmatrix} 1 & 7 \end{bmatrix} \quad \textit{(reduced)}.$$

$$= 1/100\begin{bmatrix} 1 & 7 \\ 1 & 7 \end{bmatrix}$$

• The Condition Number and Norm

The condition number $c(K) = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}$ for symmetric positive definite K.
It measures the "sensitivity" of the linear system $Ku = f$.

$\triangle f$: measurement error, roundoff etc.
$Ku = f \rightsquigarrow K(u + \triangle u) = f + \triangle f.$
The error equation: $K\triangle u = \triangle f \Rightarrow \triangle u = K^{-1}\triangle f$

$$\|\triangle u\| \le \lambda_{\text{max}}(K^{-1})\|\triangle f\| = \frac{1}{\lambda_{\text{min}}(K)}\|\triangle f\|,$$

since it is maximized when $\triangle f = q_{\text{min}}$. i.e. $K^{-1}q_{\text{min}} = \frac{1}{\lambda_{\text{min}}(K)}q_{\text{min}}$.
The λ_{min} indicates how close K to a singular matrix.
With $c >> 1$, $\lambda_{\text{min}}(cK) = c\lambda_{\text{min}}(K)$, it is far away from singular.
But if we multiply K by 1000 for examle, then u and $\triangle u$ should be divided by 1000. That rescaling to make K less singular and λ_{min} larger cannot change the reality of the problem.
The relative error $\frac{\|\triangle u\|}{\|u\|}$ stays the same.

$$\|\triangle u\| \le \frac{\|\triangle f\|}{\lambda_{\text{min}}(K)}.$$

$$\|f\| = \|Ku\| \le \lambda_{\text{max}}(K)\|u\| \Rightarrow \frac{1}{\|u\|} \le \frac{\lambda_{\text{max}}(K)}{\|f\|}.$$

$$\boxed{\frac{\|\triangle u\|}{\|u\|} \leq \frac{\lambda_{\max}(K)}{\lambda_{\min}(K)} \frac{\|\triangle f\|}{\|f\|}.}$$

Condition number for symmetric positive definite K	$c(K) = \dfrac{\lambda_{\max}(K)}{\lambda_{\min}(K)}$

Definition 19 *Matrix norm (induced)*

$$\boxed{\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|}$$

Remark *The norm of a matrix measures the maximum stretching the matrix does to any vector.*

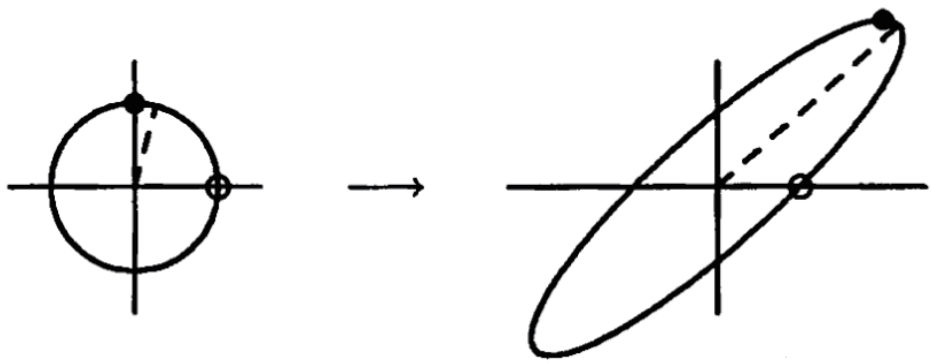

Figure 12: Matrix norm

Definition 20 *Condition number*

$$\boxed{c(A) = \|A\|\|A^{-1}\|}$$

$$\frac{\|Ax\|}{\|x\|} \leq \|A\| \text{ for all } x \neq 0 \Rightarrow \|Ax\| \leq \|A\|\|x\|.$$

$$\|AB\| \leq \|A\|\|B\| \quad \text{and} \quad \|A+B\| \leq \|A\| + \|B\|.$$

$$\|A\|^2 = \max_{x\neq 0} \frac{\|Ax\|^2}{\|x\|^2} = \max_{x\neq 0} \underbrace{\frac{x^T A^T A x}{x^T x}}_{\text{Raleigh quotient}} = \lambda_{\text{max}}(A^T A) = \sigma^2_{\text{max}}.$$

$$c(A) = \|A\|\|A^{-1}\| = \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}}.$$

$$1 = \|I\| = \|AA^{-1}\| \le \|A\|\|A^{-1}\|.$$

$$\lambda_{\text{max}}(A) \le \sigma_{\text{max}}(A).$$

Example 21 $A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 2 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = 0, \qquad \lambda_{max} = 0.$$

$$A^T A = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}, \qquad \sigma_{max} = 2 = \|A\|.$$