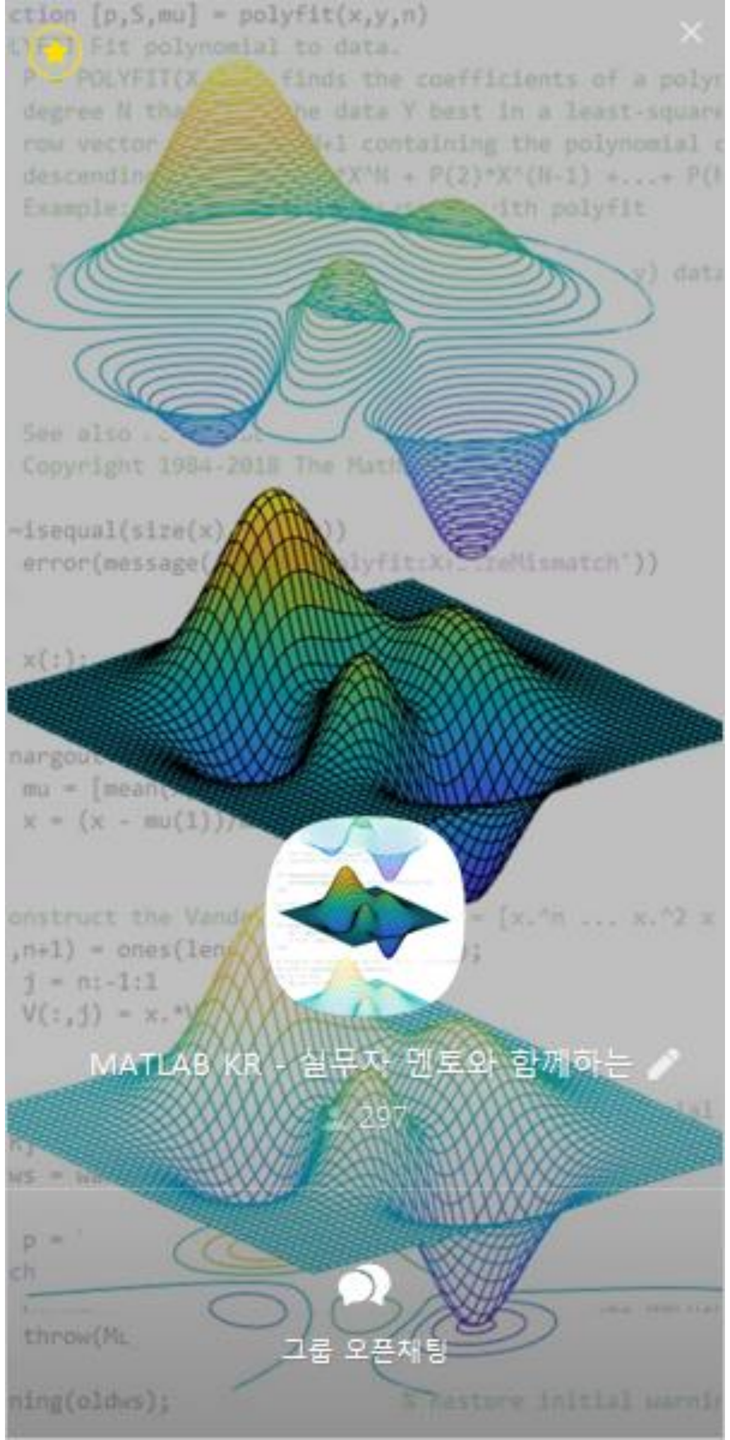
Lecture 2
Linear transformations
and Diagonalization

refs.
Contemporary Linear Algebra, H Anton, RC Busby



카카오톡 단톡방 MATLAB KR

https://open.kakao.com/o/gKBr7Zde

1. Linear transformation

1.1 Standard matrix for a linear transformation

So, what is a matrix?

- Five perspectives on viewing matrices
    - Collection of data (numeric)
    - Collection of vectors (geometric)
    - A system of linear equations (algebraic)
    - **A linear transformation (operational)**
    - A tangent space of a function (differential)

Matrix as a linear transformation

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$$

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 1x_2 \\ 0x_1 + 2x_2 \\ 1x_1 + 1x_2 \end{bmatrix}$$

- The matrix A represents a function

$$T_A : \mathbb{R}^2 \to \mathbb{R}^3$$

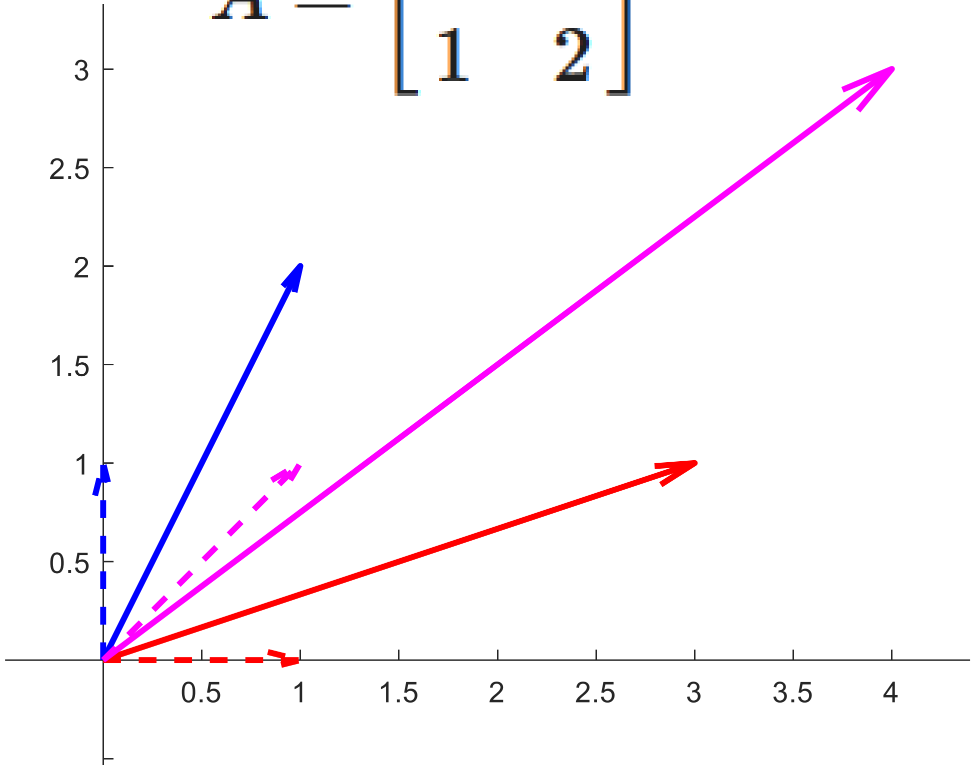
- T_A is *linear*, i.e.,

$$T_A(u + v) = T_A(u) + T_A(v)$$

$$T_A(cu) = cT_A(u)$$

Matrix as a linear transformation

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$



- In fact, we *do not need* a matrix.
- All we need is the *images of basis vectors*.
- Standard basis of $\mathbb{R}^2$

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Standard matrix

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{aligned} T(\mathrm{x}) &= x_1 T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + x_2 T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= [T(\mathrm{e}_1) \quad T(\mathrm{e}_2)] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \qquad A \qquad\qquad x \end{aligned}$$

- Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is represented by an $m \times n$ matrix.
- The inverse also holds.

- Standard matrix for T, denoted by $[T]$ ➔ $T(x) = [T]x$

Examples of linear transformations

- Stretch / compression
- Shear
- Rotation
- Reflection

- lab: example_matrix_is_a_linear_transformation.m

- Q. Is translation a *linear* transformation? (cf. Affine transformation)

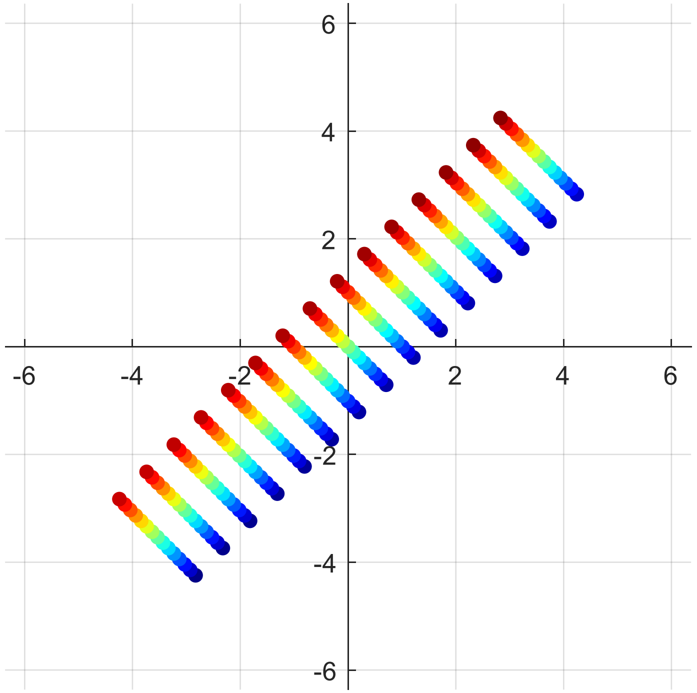
source: 공돌이의 수학 정리 노트

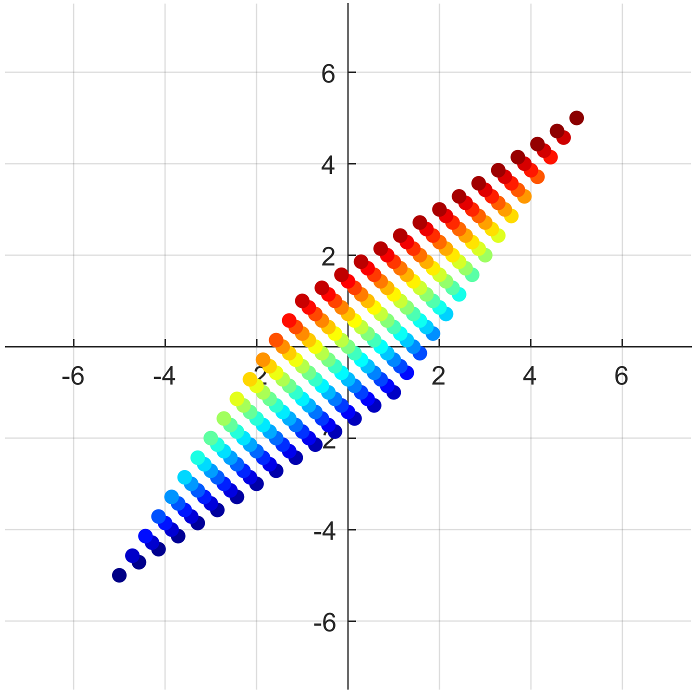
Invertibility of a square matrix

- An $n \times n$ matrix A is invertible if and only if
  - $Ax = 0$ has only a trivial solution, i.e., $\ker(T_A) = \{0\}$.
  - $\det(A) \neq 0$
  - A set of column (row) vectors of A forms a basis of $\mathbb{R}^n$, i.e., $\mathrm{Im}(T) = \mathbb{R}^n$.
  - $\mathrm{rank}(A) = n$, $\mathrm{nullity}(A) = 0$
  - T_A is 1-1 and onto.

- Q. Can an $m \times n$ matrix represent a bijective function?

Direction of stretching

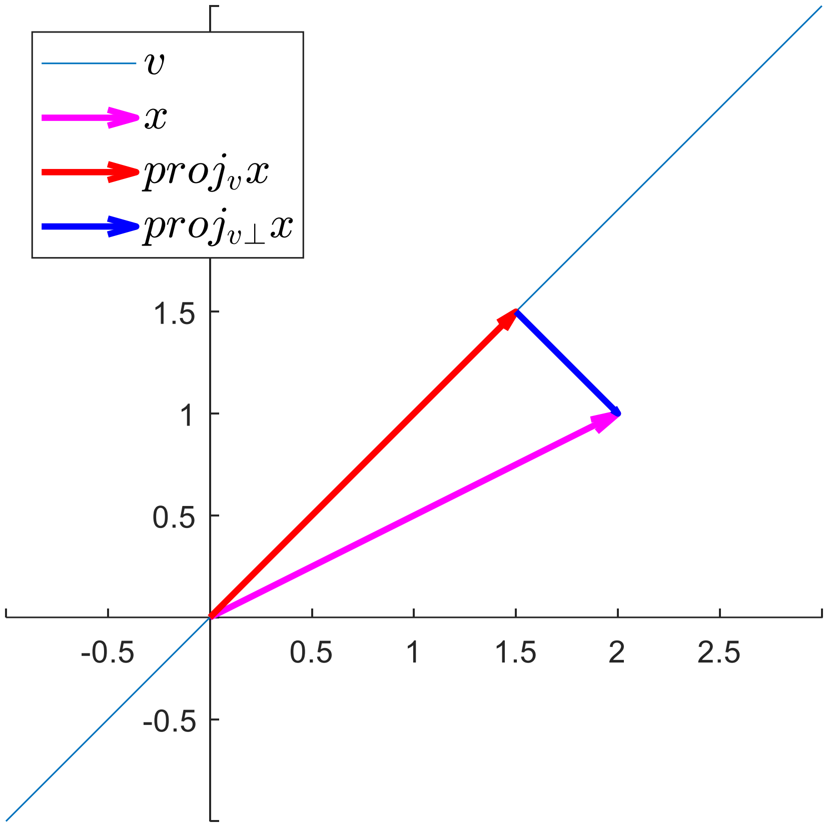




- lab: example_matrix_is_a_linear_transformation.m

1.2 Projection and least square

Projection formula



- Projection of x onto v

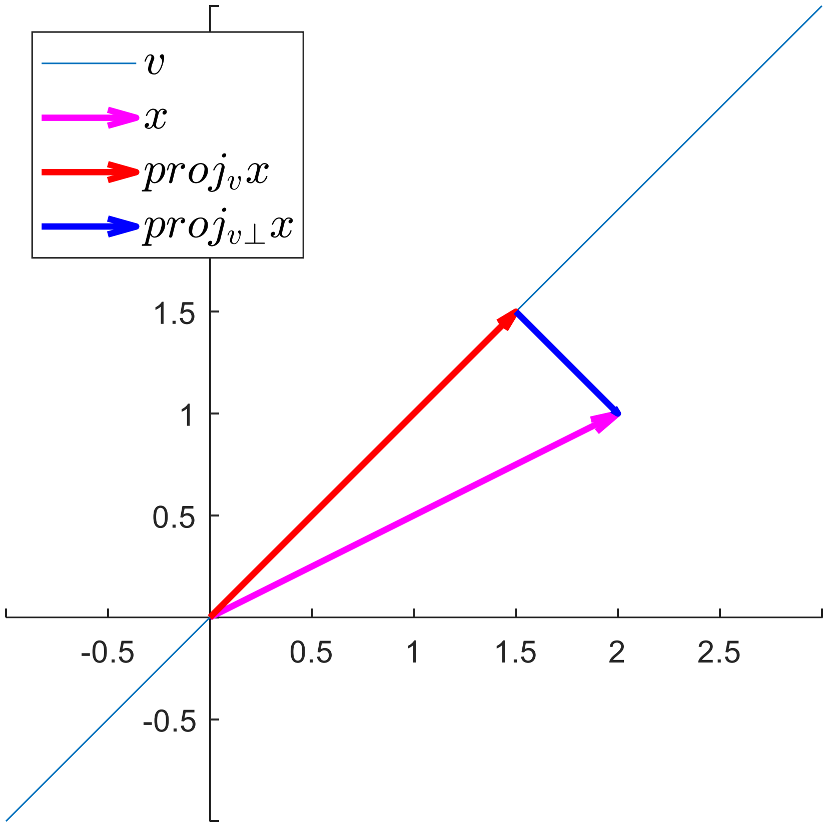
$$\mathrm{proj}_v x = \frac{x \cdot v}{\|v\|^2} v$$

$$= (x \cdot v)v \qquad \text{if } \|v\| = 1$$

- Every vector x in $\mathbb{R}^n$ is *uniquely* expressed as

$$x = \mathrm{proj}_v x + \mathrm{proj}_{v^\perp} x$$

Projection operator



- Projection operator onto span$\{v\}$

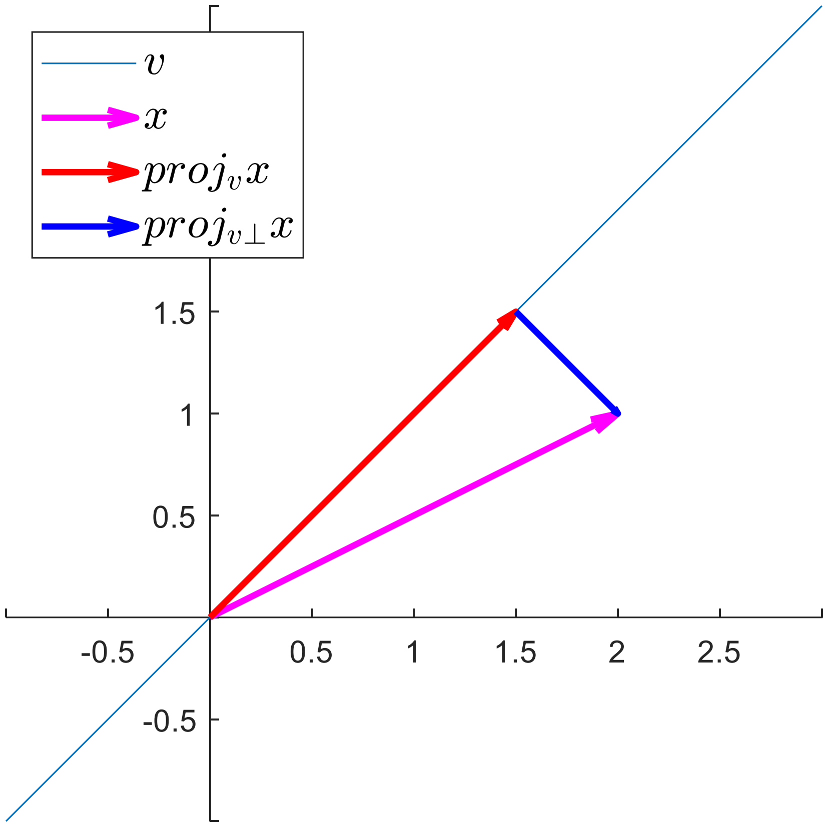
$$T(x) = \mathrm{proj}_v x = \left(\frac{1}{v^T v} v v^T\right) x$$

- Standard matrix for T

$$P = \frac{1}{v^T v} v v^T \qquad (= v v^T \text{ if } \|v\| = 1)$$

- Q. $\mathrm{rank}(P) = ?$

Projection onto a *subspace*



- Projection of $x \in \mathbb{R}^n$ onto $W \leq \mathbb{R}^n$

$$\mathrm{proj}_W x = M(M^T M)^{-1} M^T x$$

- where the column vectors of M form a basis of W.

- lab: example_projection_operator.m

Least square and projection operation

Linear least squares [edit]

Main article: Linear least squares

A regression model is a linear one when the model comprises a linear combination of the parameters, i.e.,

$$f(x, \boldsymbol{\beta}) = \sum_{j=1}^{m} \beta_j \phi_j(x),$$

where the function ϕ_j is a function of x.[12]

Letting $X_{ij} = \phi_j(x_i)$ and putting the independent and dependent variables in matrices X and Y, respectively, we can compute the least squares in the following way. Note that D is the set of all data.[12][13]

$$L(D, \boldsymbol{\beta}) = \|Y - X\boldsymbol{\beta}\|^2 = (Y - X\boldsymbol{\beta})^\mathsf{T}(Y - X\boldsymbol{\beta}) = Y^\mathsf{T}Y - Y^\mathsf{T}X\boldsymbol{\beta} - \boldsymbol{\beta}^\mathsf{T}X^\mathsf{T}Y + \boldsymbol{\beta}^\mathsf{T}X^\mathsf{T}X\boldsymbol{\beta}$$
$$= Y^\mathsf{T}Y - X^\mathsf{T}Y\boldsymbol{\beta} - X^\mathsf{T}Y\boldsymbol{\beta} + X^\mathsf{T}X\boldsymbol{\beta}^2$$

The gradient of the loss is:

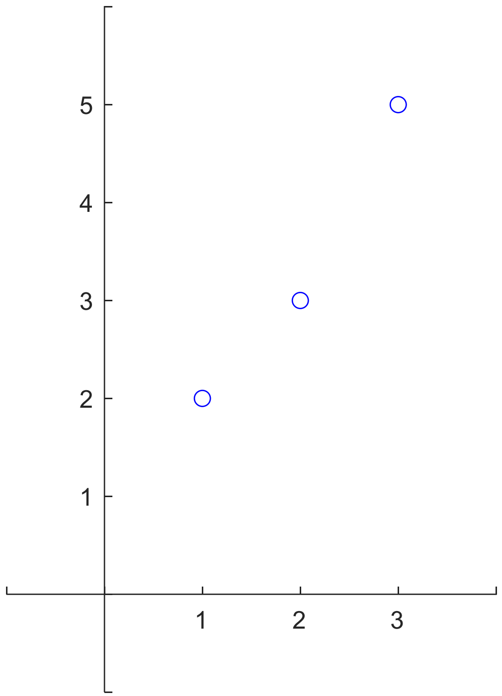
$$\frac{\partial L(D, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial \left(Y^\mathsf{T}Y - X^\mathsf{T}Y\boldsymbol{\beta} - X^\mathsf{T}Y\boldsymbol{\beta} + X^\mathsf{T}X\boldsymbol{\beta}^2\right)}{\partial \boldsymbol{\beta}} = -2X^\mathsf{T}Y + 2X^\mathsf{T}X\boldsymbol{\beta}$$

Setting the gradient of the loss to zero and solving for $\boldsymbol{\beta}$, we get:[13][12]

$$-2X^\mathsf{T}Y + 2X^\mathsf{T}X\boldsymbol{\beta} = 0 \Rightarrow X^\mathsf{T}Y = X^\mathsf{T}X\boldsymbol{\beta}$$
$$\boxed{\hat{\boldsymbol{\beta}} = (X^\mathsf{T}X)^{-1}X^\mathsf{T}Y}$$

$$\boxed{\mathrm{proj}_W x = M(M^T M)^{-1} M^T x}$$

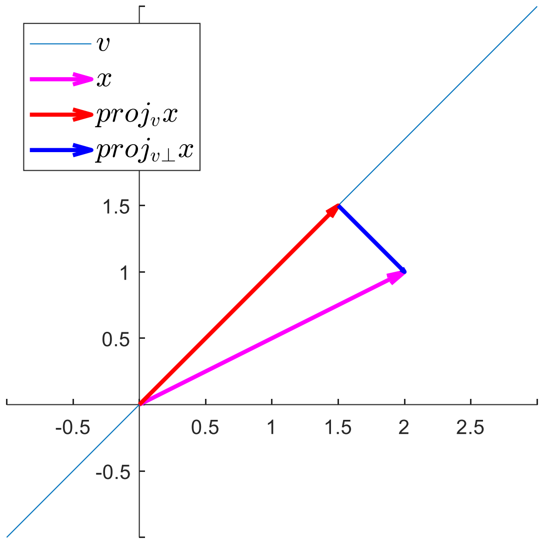
Ordinary Least Squares



$$c_0 + c_1 x_1 + e_1 = y_1$$
$$c_0 + c_1 x_2 + e_2 = y_2$$
$$c_0 + c_1 x_3 + e_3 = y_3$$

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$A \quad x \quad + \quad e \quad = \quad y$$



- Least square method

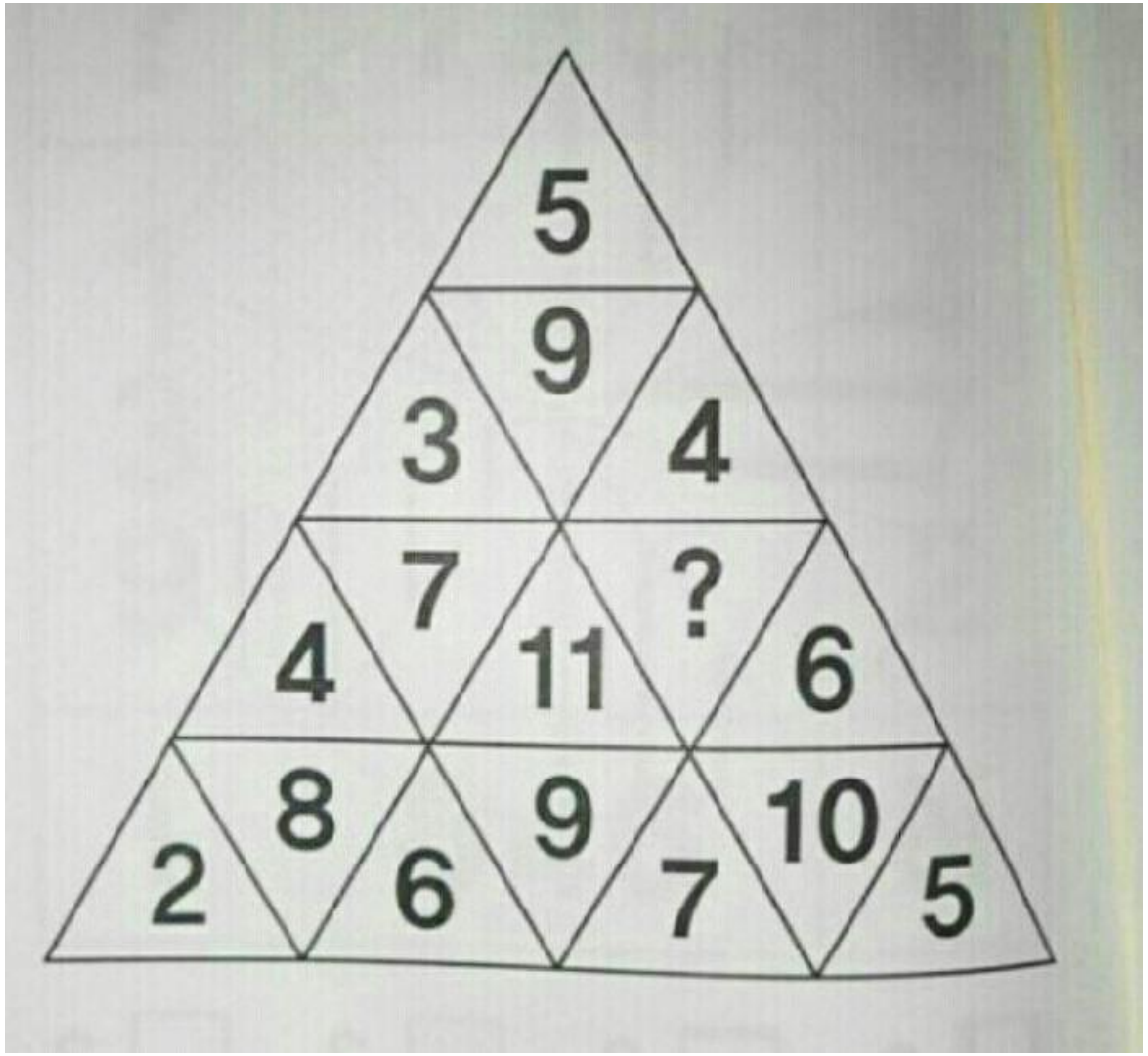
  ≡ To find x and $\hat{y}$ that minimize ||e||

  ≡ Projecting y onto $\mathrm{col}(A)$ which is $\hat{y}$

$$\boxed{\mathrm{proj}_{\mathrm{col}(A)} y = A(A^T A)^{-1} A^T y}$$

- lab: example_least_square_is_projection.m

Linear regression example(?)



- lab: example_triangle_game.m

Orthonormal basis

- $v \in \mathbb{R}^n, \|v\| = 1$

➔ Projection operator onto span$\{v\}$

$$T(x) = (vv^T)x$$

- Standard matrix for T

- M is an orthonormal basis of $W \leq \mathbb{R}^n$

➔ Projection operator onto col(M)

$$T(x) = M(M^TM)^{-1}M^Tx$$

$$= MM^Tx$$

Standard matrix for T

Why do we care orthonormal bases?

- [Thm] If $\{v_1, \dots, v_k\}$ is an orthonormal basis for $W \leq \mathbb{R}^n$, then the orthogonal projectin of a vector $x \in \mathbb{R}^n$ onto W is expressed as

$$\mathrm{proj}_W x = (x \cdot v_1)v_1 + (x \cdot v_2)v_2 + \cdots + (x \cdot v_k)v_k$$

- [Thm] If $\{v_1, \dots, v_k\}$ is an orthonormal basis for $W \leq \mathbb{R}^n$, then any $w \in W$ is expressed as

$$w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \cdots + (w \cdot v_k)v_k$$

Gram-Schmidt and QR decomposition

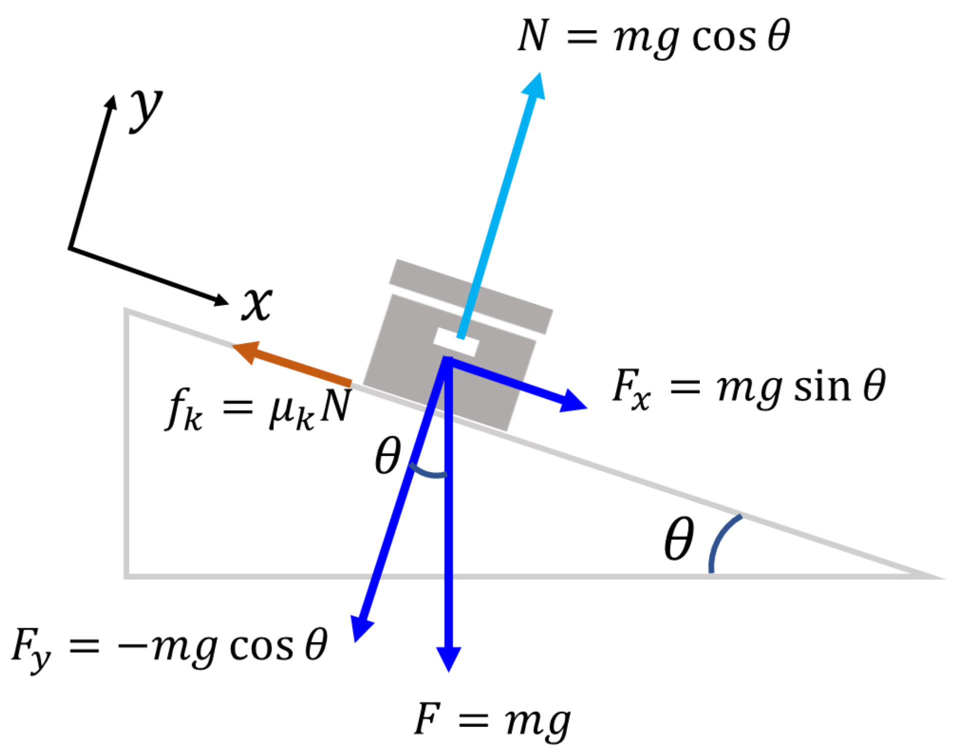
- From a basis $\{w_1, \ldots, w_k\}$ for $W \leq \mathbb{R}^n$, an orthonormal basis of W can be constructed as follows:
  - 1. Normalize w_1 to u_1 such that $|u_1| = 1$.
  - 2. Calculate $v_2 = w_2 - \mathrm{proj}_{u_1} w_2$ and normalize to u_2.
  - 3. Calculate $v_3 = w_3 - \mathrm{proj}_{\mathrm{span}\{u_1, u_2\}} w_3$ and normalize to u_3.
  - 4. Repeat to v_k and u_k.
- Rmk. $\mathrm{span}\{w_1, \ldots, w_l\} = \mathrm{span}\{u_1, \ldots, u_l\}$ for any $l \leq k$.

- lab: example_Gram_Schmidt_and_qr_decomposition.m

2. Coordinate and change of basis

2.1 Vectors and coordinates

Making the familiar *Unfamiliar*



$$\begin{bmatrix} mg\sin\theta \\ -mg\cos\theta \end{bmatrix} \quad \text{vs} \quad \begin{bmatrix} 0 \\ -mg \end{bmatrix}$$

- What is a vector?
  - An arrow?
  - A tuple of numbers?
  - Something with magnitude and direction?
- Vector ≠ Coordinate
- Standard basis is not that *special*.

Coordinate of a vector

- [Def] If $B = \{v_1, \dots, v_k\}$ is an ordered basis for $W \leq \mathbb{R}^n$, and if

$$w = a_1 v_1 + \cdots + a_k v_k$$

- then we call $a_1, \dots, a_k$ the *coordinate* of w with respect to B, and

$$[w]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}$$

- the coordinate matrix for w with respect to B.

Why do we care orthonormal bases?

- [Thm] If B is an orthonormal basis for a k-dim. $W \leq \mathbb{R}^n$, and if u, v, and w are vectors in W such that

$$[u]_B = \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix}, [v]_B = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}, [w]_B = \begin{bmatrix} w_1 \\ \vdots \\ w_k \end{bmatrix}$$

- then
  - (a) $\|w\| = \sqrt{w_1^2 + w_2^2 + \cdots + w_k^2} = \|[w]_B\|$
  - (b) $u \cdot v = u_1 v_1 + u_2 v_2 + \cdots u_k v_k = [u]_B \cdot [v]_B$

Change of basis

- How to change from $[w]_B$ to $[w]_{B'}$?
- [Thm] Given $w \in \mathbb{R}^n$, two bases of $\mathbb{R}^n$ $B = \{v_1, \dots, v_n\}$ and $B' = \{v_1', \dots, v_n'\}$,

$$[w]_{B'} = P_{B \to B'} [w]_B$$

- where $P_{B \to B'}$ is the coordinate change matrix from B to B'

$$P_{B \to B'} = \left[[v_1]_{B'}, \dots, [v_n]_{B'} \right]$$

- Rmk. $P_{B' \to B} = (P_{B \to B'})^{-1}$
- Q. $P_{B \to B} =?$

Change of basis

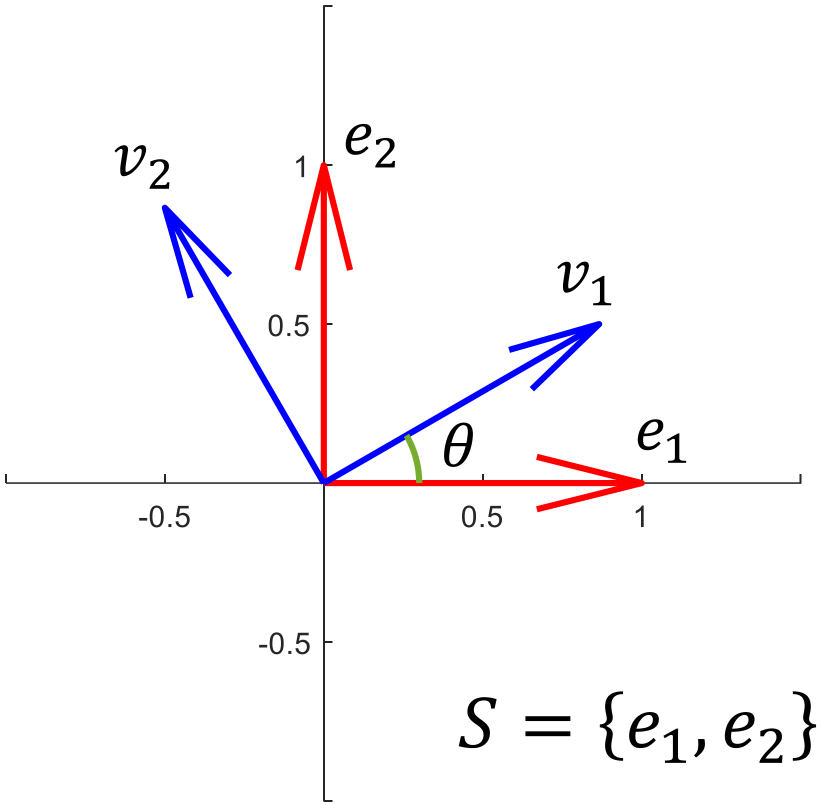
- If S is the standard basis of $\mathbb{R}^n$, then $P_{B\to S} = \big[[v_1]_S, \dots, [v_n]_S\big] = [v_1, \dots, v_n]$.

- Rmk. Change of basis from B to B' = from B to S, and then from S to B'.

$$P_{B\to B'} = P_{S\to B'}P_{B\to S} = (P_{B'\to S})^{-1}P_{B\to S} = \big[[v_1']_S, \dots, [v_n']_S\big]^{-1}\big[[v_1]_S, \dots, [v_n]_S\big]$$

- [Thm] If B and B' are orthonormal bases for $\mathbb{R}^n$, then the coordinate change matrices $P_{B\to B'}$ and $P_{B'\to B}$ are orthogonal.

Change of basis – example

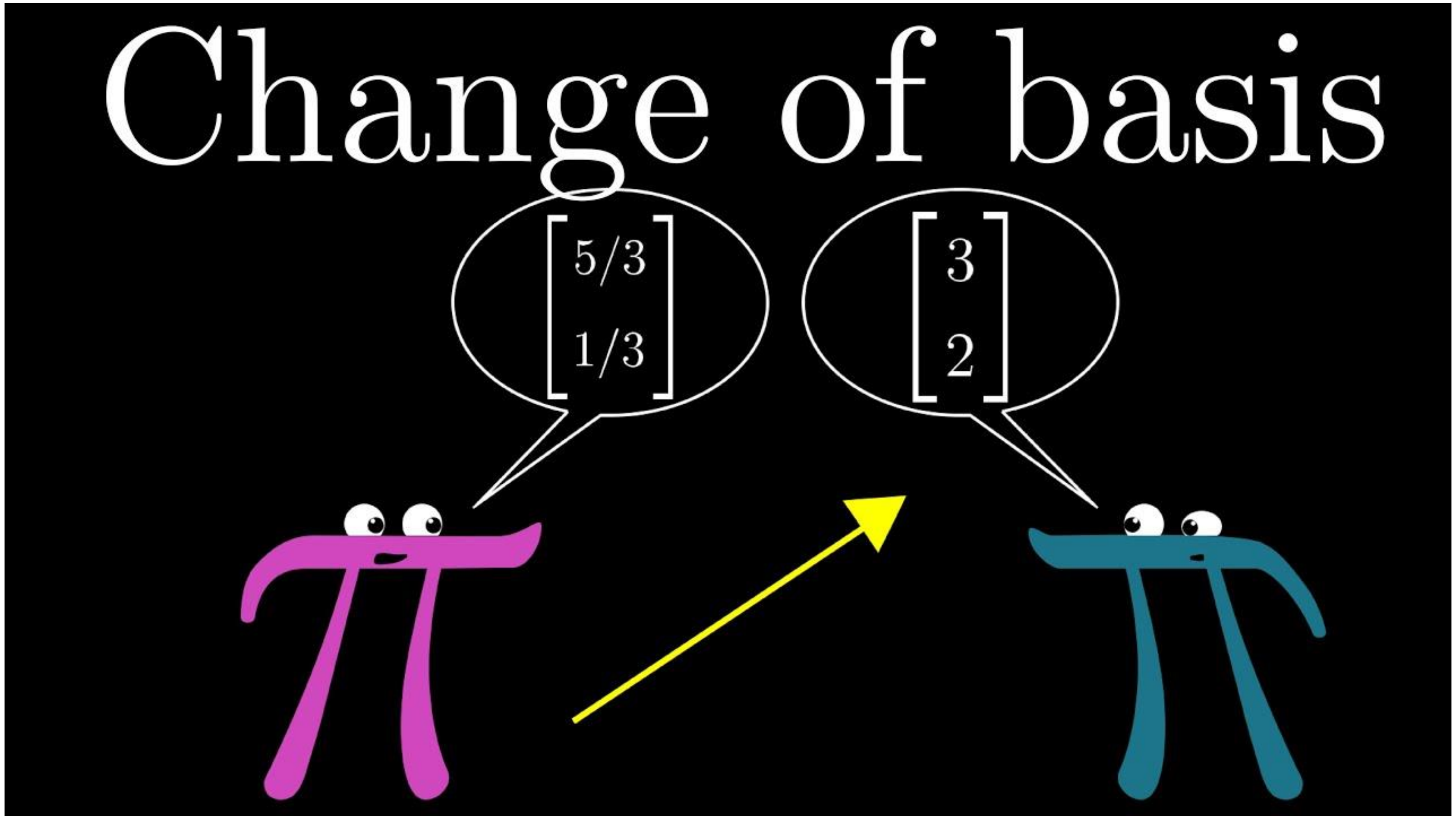


$$S = \{e_1, e_2\}$$

$$B = \{v_1, v_2\}$$

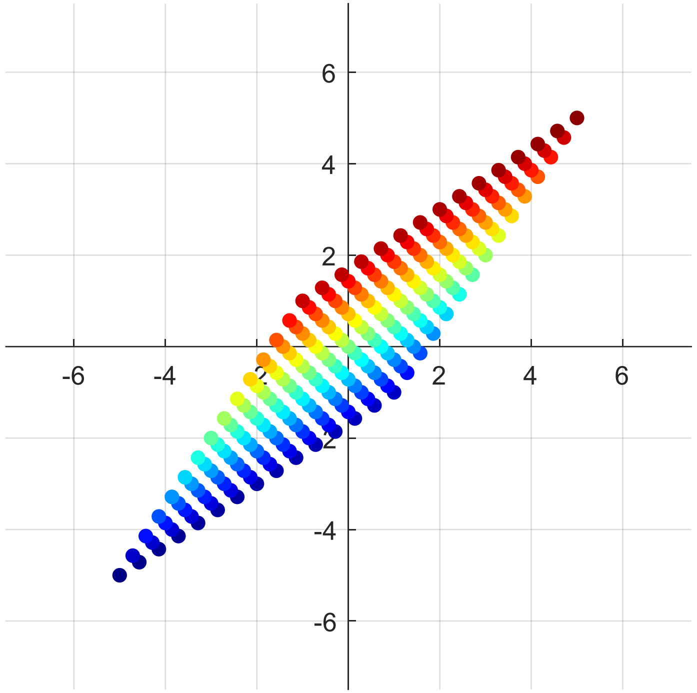
- How to interpret $V = [v_1, v_2]$?
- $V = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- 1) Rotation matrix
- 2) Coordinate change matrix $P_{B \to S}$

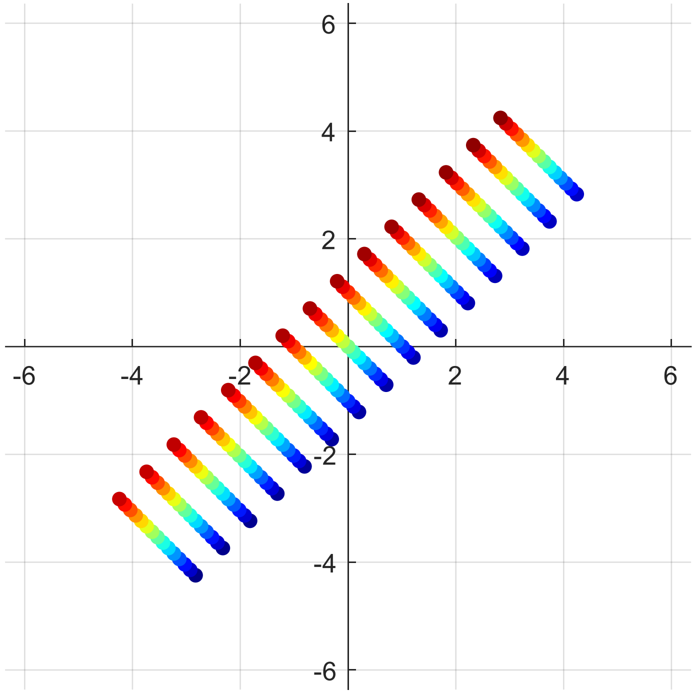
Change of basis



2.2 Matrix representation of linear transformations

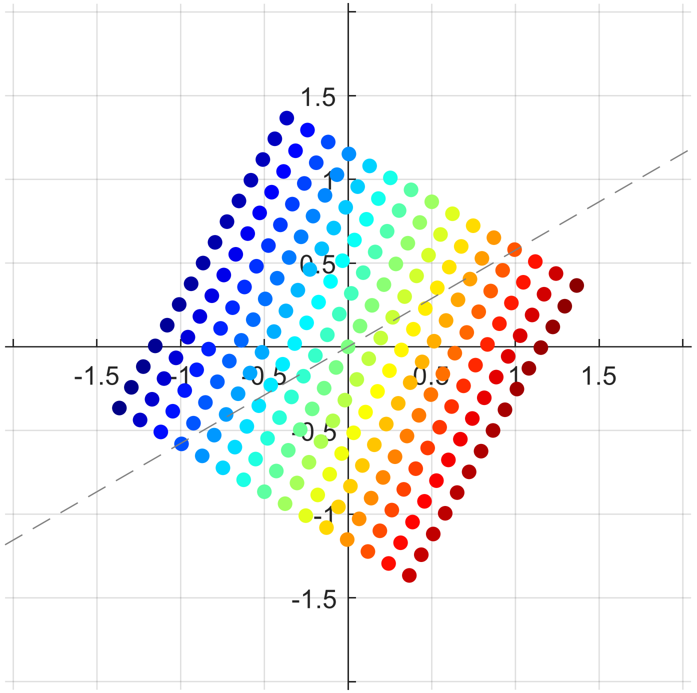
Do linear transformations depend on bases?

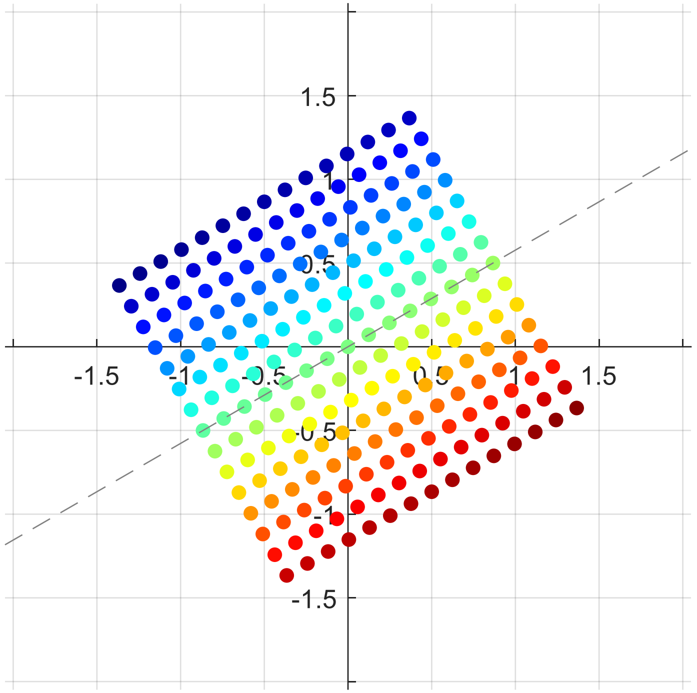




$$\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}_S \qquad \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}_B$$

Do linear transformations depend on bases?





$$\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}_S$$

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}_B$$

Do linear transformations depend on bases?

$$x \xrightarrow{\quad T \quad} T(x)$$

$$[x]_B \xrightarrow{\quad A \quad} A[x]_B$$

- The coordinate of x depends on the basis.
- T transforms a vector to another vector.
- What does the matrix A transform?
  - Vectors?
  - Coordinates of vectors?
- Is the matrix A for T unique?

Do linear transformations depend on bases?

- A matrix represents a linear transformation with respect to a *specific* basis.
- *Matrix representation* of a linear transformation depends on the *choice of basis*.
- Standard matrix: Matrix representation of a linear transformation with respect to the *standard* basis.

$$
\begin{aligned}
T(\mathrm{x}) &= x_1 T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + x_2 T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) \\
&= [T(\mathrm{e}_1) \quad T(\mathrm{e}_2)] \begin{bmatrix}x_1\\x_2\end{bmatrix} \\
&= \qquad A \qquad\qquad x
\end{aligned}
$$

Matrix representation of linear transformations

- [Thm] Let T: $\mathbb{R}^n \to \mathbb{R}^n$ be a linear operator, $B = \{v_1, \dots, v_n\}$ be a basis for $\mathbb{R}^n$. Then $A = \left[[T(v_1)]_B, \dots, [T(v_n)]_B\right]$ is called the matrix for T with respect to the basis B, denoted by $[T]_B$, and satisfies

$$[T(x)]_B = A[x]_B = [T]_B[x]_B$$

- Q. What is a vector? (hint: not a tuple)
- Q. What is a matrix? (hint: not a rectangular arrangement of numbers)
- Q. What is a linear transformation? (hint: What is the linearity?)
- Q. How to get $[T]_B$ from $[T]_{B'}$?

Matrix representation of linear transformations

$$
\begin{tikzcd}
{[x]_S} \arrow[r, "{[T]_S = A}"] \arrow[d, "P^{-1}"'] & {[T(x)]_S} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}
$$

$$x \xrightarrow{T} T(x)$$

$$[x]_B \xrightarrow{[T]_B} [T(x)]_B = [T]_B[x]_B$$

- Recall that the coordinate change matrix from an ordered basis $B = \{v_1, \dots, v_n\}$ to the standard basis S is

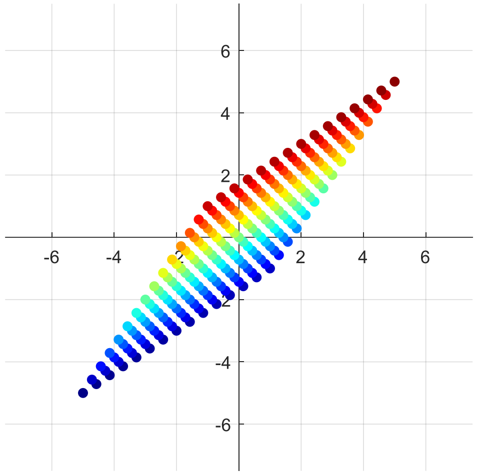
$$P_{B\to S} = [v_1, \dots, v_n]$$

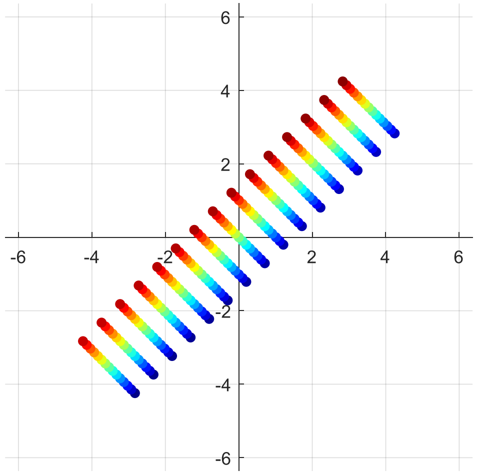
- Thus, the followings hold.

$$[T]_S = P[T]_B P^{-1}$$

$$[T]_B = P^{-1}[T]_S P$$

Matrix representation of linear transformations





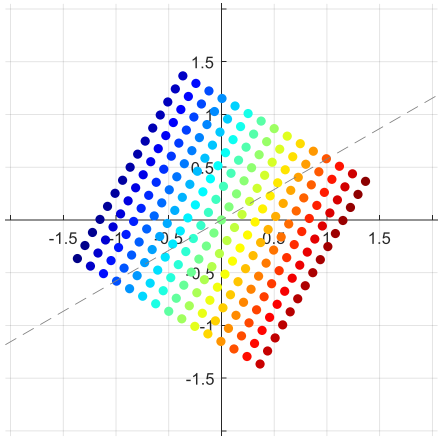
- Example

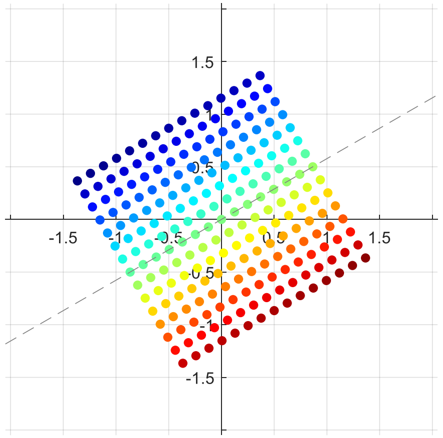
$$S = \{e_1, e_2\}, B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

$$P_{B \to S} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$[T]_S = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

Matrix representation of linear transformations





- Example

$$S = \{e_1, e_2\}, B = \left\{ \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \right\}$$

$$P_{B\to S} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$[T]_S = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

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

Matrix representation of linear transformations

$$\begin{tikzcd}[column sep=huge, row sep=huge]
{[x]_{B'}} \arrow[r, "{[T]_{B'}}"] \arrow[d, "P^{-1}"'] & {[T(x)]_{B'}} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}$$

- For genelarlization, given a coordinate change matrix P from B to B', the followings hold.

$$[T]_{B'} = P[T]_B P^{-1}$$

$$[T]_B = P^{-1}[T]_{B'} P$$

Matrix representation of linear transformations

- Bases of domain and codomain of a linear transformation may be *different*.
- [Thm] Let $T\colon\ \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, $B = \{v_1, \dots, v_n\}$ be a basis for $\mathbb{R}^n$, $B' = \{u_1, \dots, u_m\}$ be a basis for $\mathbb{R}^m$. Then

$$A = \left[[T(v_1)]_{B'}, \dots, [T(v_n)]_{B'}\right]$$

- is the matrix for T with respect to the bases B and B', denoted by $[T]_B^{B'}$, and satisfies

$$[T(x)]_{B'} = A[x]_B = [T]_B^{B'}[x]_B$$

Matrix representation of linear transformations

coordinate change from B_1 to B_2, denoted by $[\mathrm{Id}]_{B_1}^{B_2}$

$$\begin{tikzcd}
{[x]_{B_1}} \arrow[r, "{[T]_{B_1}^{B_1'}}"] \arrow[d, "V^{-1}"'] & {[T(x)]_{B_1'}} \\
{[x]_{B_2}} \arrow[r, "{[T]_{B_2}^{B_2'}}"'] & {[T(x)]_{B_2'}} \arrow[u, "U"']
\end{tikzcd}$$

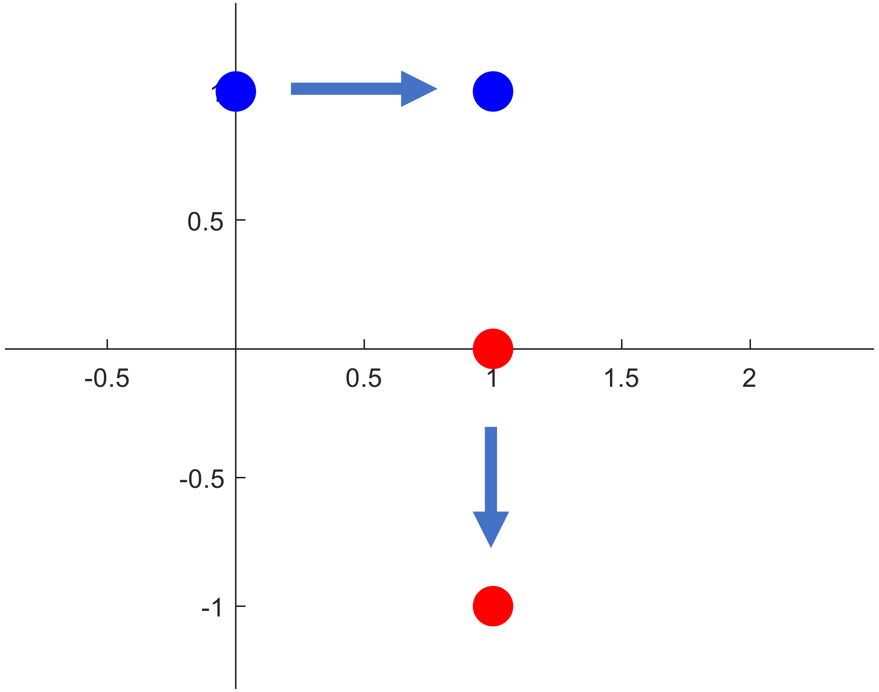
coordinate change from B_2' to B_1', denoted by $[\mathrm{Id}]_{B_2'}^{B_1'}$

$$[T]_{B_1}^{B_1'} = U[T]_{B_2}^{B_2'}V^{-1} = [\mathrm{Id}]_{B_2'}^{B_1'}[T]_{B_2}^{B_2'}[\mathrm{Id}]_{B_1}^{B_2}$$

$$[T]_{B_1}^{B_1'}[x]_{B_1} = [\mathrm{Id}]_{B_2'}^{B_1'}[T]_{B_2}^{B_2'}[\mathrm{Id}]_{B_1}^{B_2}[x]_{B_1}$$

Matrix representation of linear transformations

- Rmk. Since the matrix representation depends on the both bases of domain and codomain, an identity matrix may **not** represent an identity transformation.



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

$$[T]_S^B = [\mathrm{Id}]_S^B [T]_S$$

3. Diagonalization

Some quote

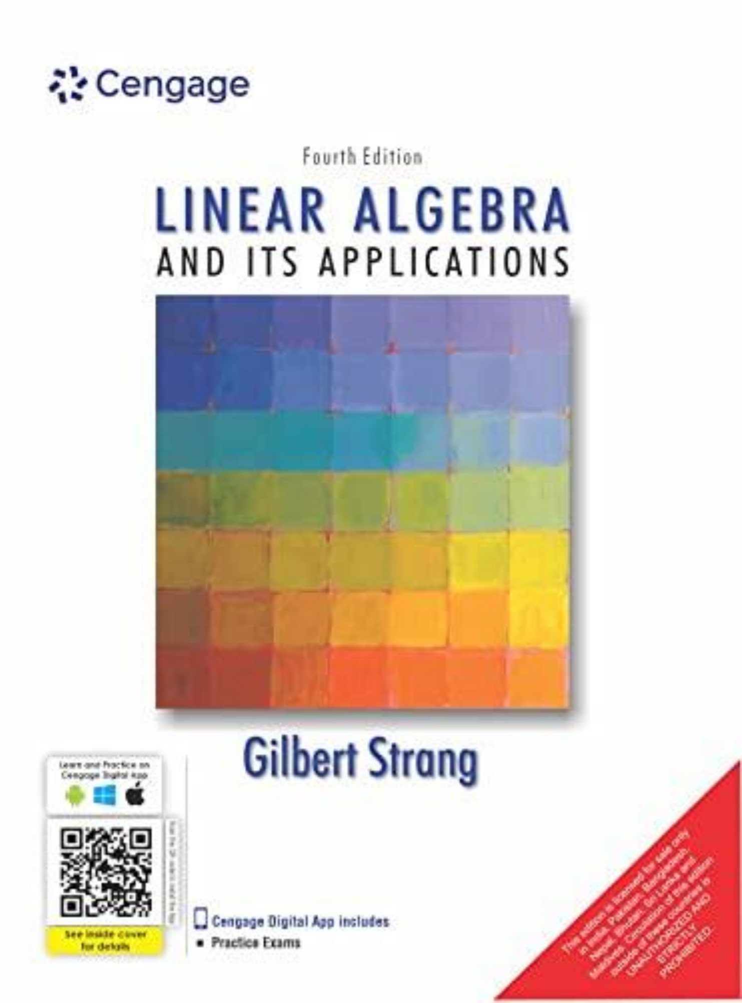
- "This chapter begins the *"second half"* of linear algebra.

  The first half was about $Ax = b$.

  The new problem $Ax = \lambda x$ will still be solved

  making it *diagonal* if possible."

  - Linear Algebra and Its Applications, Gilbert Strang.



3.1 Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

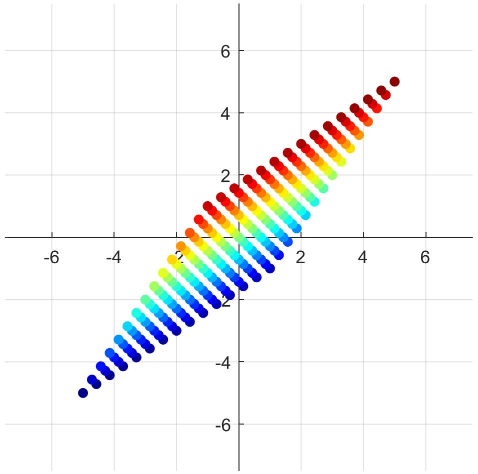
- [Def] If A is an $n \times n$ matrix, then a scalar λ is called an eigenvalue of A if there is a **nonzero** vector x such that $Ax = \lambda$. Such a vector x is called an eigenvector of A corresponding to λ.

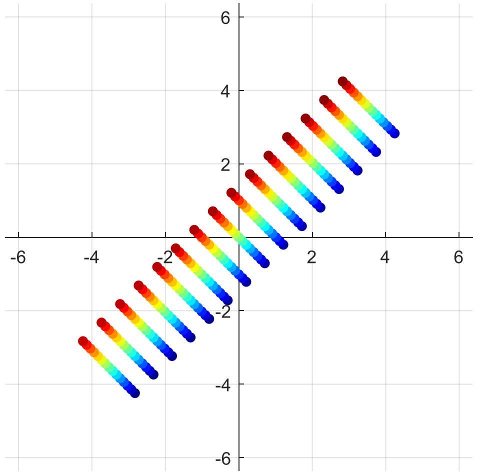
- Rmk. Eigenvalues are the solutions of $|\lambda I - A| = 0$.

- [Def] For an $n \times n$ matrix A, $|\lambda I - A|$ is called the *characteristic polynomial* of A.

- Q. Does a rotation matrix have eigenvalues?

Eigenvalues and Eigenvectors





- Example

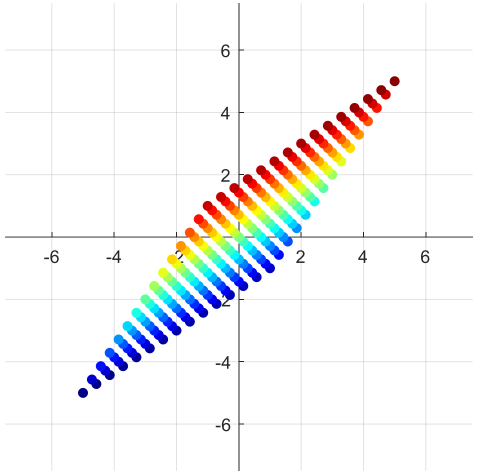
$$S = \{e_1, e_2\}, B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

$$P_{B \to S} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$[T]_S = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

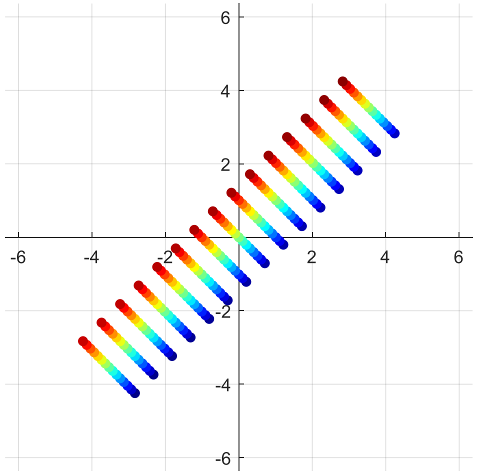
- What are the eigenvalues and eigenvectors of $[T]_S$?

Eigenvalues and Eigenvectors



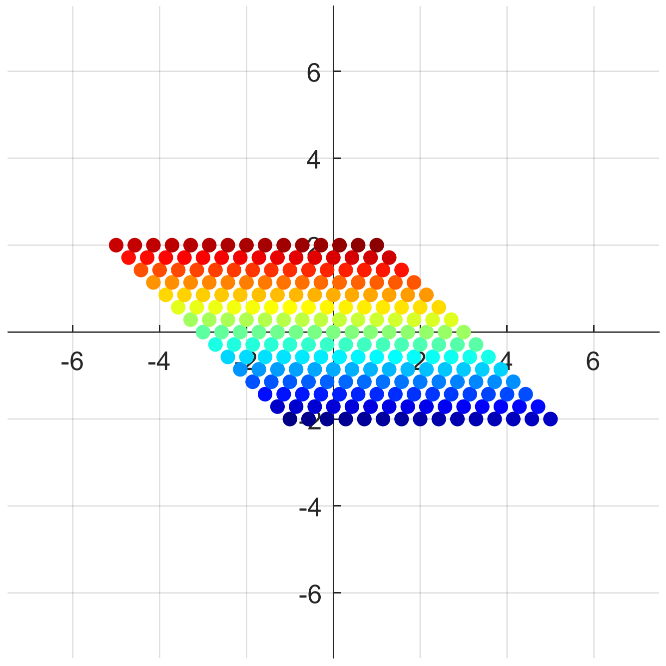
$$S = \{e_1, e_2\} \qquad \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

$$\begin{tikzcd}[column sep=huge, row sep=huge]
{[x]_S} \arrow[r, "{[T]_S}"] \arrow[d, "P^{-1}"'] & {[T(x)]_S} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}$$



$$B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \qquad \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$$

Eigenvalues and Eigenvectors



$$S = \{e_1, e_2\}, B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

$$P_{B \to S} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

$$[T]_S = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1}$$

- lab: example_direction_of_eigenvectors.m

Linear independence of eigenvectors

- [Thm] If $v_1, \dots, v_k$ are eigenvectors of a matrix A that corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_k$, then the set $\{v_1, \dots, v_k\}$ is linearly independent.

- [Cor] If an $n \times n$ matrix A has n distinct eigenvalues, then the set of eigenvectors of A forms a basis of $\mathbb{R}^n$.

- Q. What if an $n \times n$ matrix A have less than n eigenvalues?

Multiplicity of eigenvalues and eigenspaces

- [Def] If the characteristic polynomial of an $n \times n$ matrix A is factored as

$$|\lambda I - A| = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_k)^{m_k}$$

- then m_i is called the **algebraic multiplicity** of the eigenvalue λ_i.

- [Def] If λ_0 is an eigenvalue of A, the solution space of

$$(\lambda_0 I - A)x = 0$$

- is called the **eigenspace** of A corresponding to λ_0.
- The dimension of this eigenspace is called the **geometric multiplicity** of λ_0.

Multiplicity of eigenvalues and eigenspaces

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

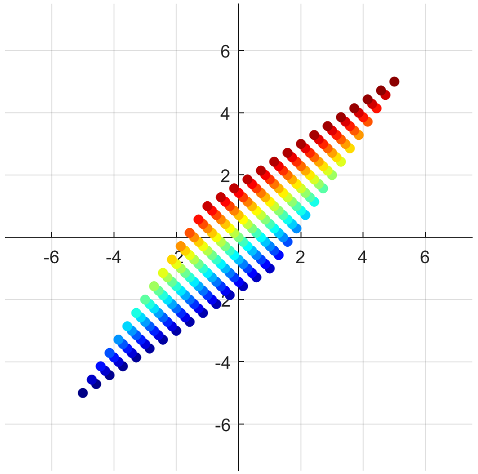
- Find the…
  - Characteristic polynomial of A
  - Eigenvalues and eigenvectors
  - Algebraic multiplicity and geometric multiplicity of each eigenvalue
  - Eigenspace for each eigenvalue
- lab: example_eigenvalue_problem.m

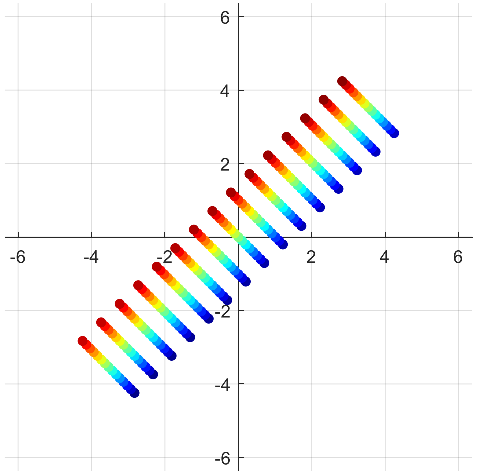
Invertibility of a square matrix

- An $n \times n$ matrix A is invertible if and only if
  - $Ax = 0$ has only a trivial solution, i.e., $\ker(T_A) = \{0\}$.
  - $\det(A) \neq 0$
  - A set of column (row) vectors of A forms a basis of $\mathbb{R}^n$, i.e., $\mathrm{Im}(T) = \mathbb{R}^n$.
  - $\mathrm{rank}(A) = n$, $\mathrm{nullity}(A) = 0$
  - T_A is 1-1 and onto.
  - 0 is not an eigenvalue of A.

3.2 Matrix similarity

What is the matrix similarity?





$$S = \{e_1, e_2\} \qquad \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

$$\begin{tikzcd}
{[x]_S} \arrow[r, "{[T]_S}"] \arrow[d, "P^{-1}"'] & {[T(x)]_S} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}$$

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \qquad \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$$

Matrix similarity

- [Def] If A and C are square matrices with the same size, then we say that C is similar to A, denote $A \sim C$, if there exists an invertible matrix P such that $C = P^{-1}AP$.

- [Thm] If A and C are two matrix representations of the same linear operator, then A is similar to C. The inverse is also true.

- [Cor] If $A \sim C$, then $C \sim A$.

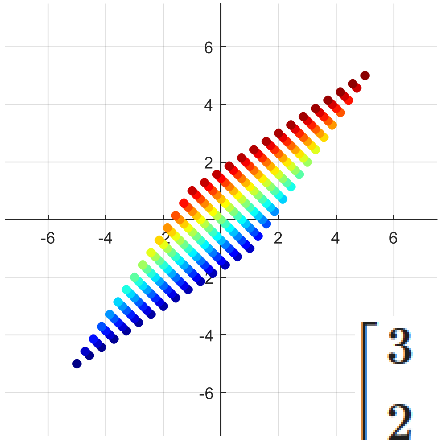
Similar matrices are *similar*

- [Thm] Similar matrices have similarity invariants such as:
  - the same determinant.
  - the same rank.
  - the same nullity.
  - the same trace.
  - the same characteristic polynomial.
  - the same eigenvalues with the same algebraic and geometric multiplicities.

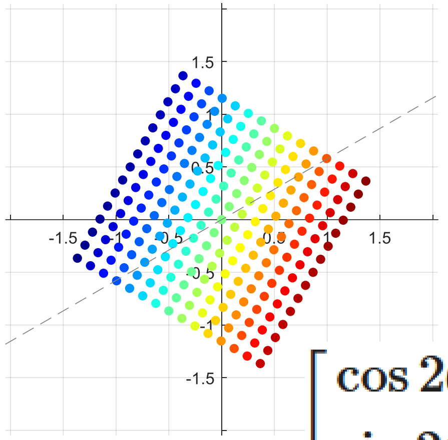
- Q. Are the properties above for a matrix or a linear operator?
- Q. If two matrices have same properties above, are they similar?
- lab: example_similarity_check.m

3.3 Eigenvalue Decomposition

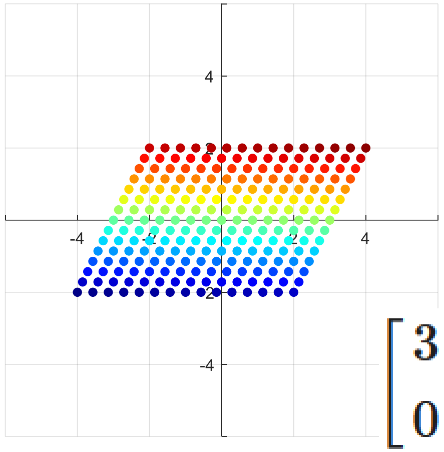
Motivation of Diagonalization



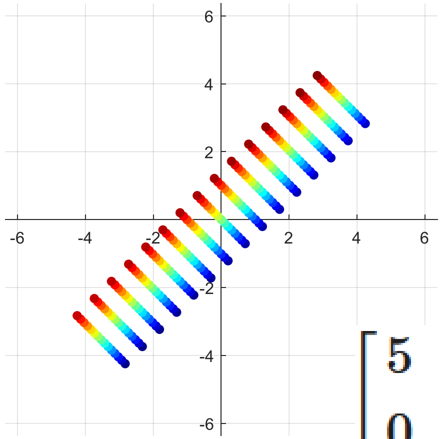
$$\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$



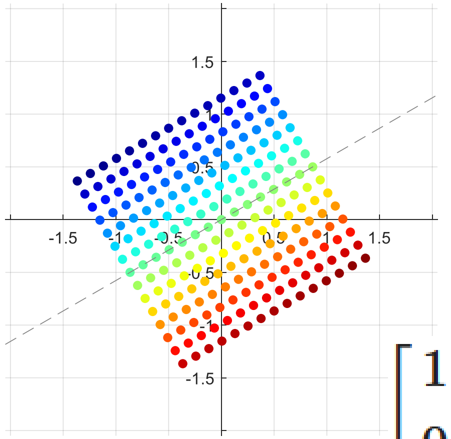
$$\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$



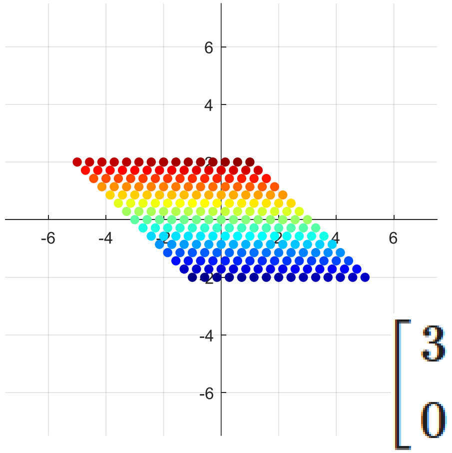
$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$



$$\begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$



$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

The Diagonalization Problem

- [Def] For a square matrix A, if there exists an invertible matrix P such that $P^{-1}AP$ is *diagonal*, then A is said to be **diagonalizable**, and P is said is to diagonalize A.

- [Thm] An $n \times n$ matrix A is diagonalizable
  $\Leftrightarrow$ A has n linearly independent eigenvectors.
  $\Leftrightarrow$ Eigenvectors of A form a basis of $\mathbb{R}^n$.

$$A = PDP^{-1}$$

Column vectors are eigenvectors. Diagonal terms are eigenvalues.

Diagonalizability of matrices

- [Thm] An $n \times n$ matrix with n distinct real eigenvalues is diagonalizable.

- [Thm] If A is an $n \times n$ matrix, then
  - (a) The geometric multiplicity of an eigenvalue of A is less than or equal to its algebraic multiplicity.
  - (b) A is diagonalizable.
    $\Leftrightarrow$ The sum of the geometric multiplicities of its eigenvalues is n.
    $\Leftrightarrow$ The geometric multiplicity of each eigenvalue of A is the same as its algebraic multiplicity.

- Rmk. Diagonalizability does not mean distinct eigenvalues.

Diagonalizability of matrices: example

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

- Is A diagonalizable?
- If so, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.
- lab: example_eigenvalue_problem.m

We still care the orthonormal bases

$$
\begin{tikzcd}
{[x]_S} \arrow[r, "{[T]_S = A}"] \arrow[d, "P^{-1}"'] & {[T(x)]_S} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}
$$

- It is almost always better to have an orthonormal basis.
- When can a square matrix A be decomposed like this:

$$A = PDP^T$$

The Orthogonal Diagonalization Problem

- [Def] For a square matrix A, if there exists an *orthogonal* matrix P such that $P^T AP$ is *diagonal*, then A is said to be **orthogonally diagonalizable**, and P is said is to orthogonally diagonalize A.

$$
\begin{tikzcd}
{[x]_S} \arrow[r, "{[T]_S = A}"] \arrow[d, "P^T"'] & {[T(x)]_S} \\
{[x]_B} \arrow[r, "{[T]_B}"'] & {[T(x)]_B} \arrow[u, "P"']
\end{tikzcd}
$$

$$A = PDP^T$$

Column vectors are eigenvectors.

Diagonal terms are eigenvalues.

Orthogonal diagonalizability of matrices

- [Thm] An $n \times n$ matrix A is orthogonally diagonalizable.
  $\Leftrightarrow$ There exists an orthonormal set of n eigenvectors of A.
  $\Leftrightarrow$ A is symmetric.

- [Thm] If a square matrix A is symmetric, then eigenvectors from different eigenspaces are orthogonal.

- Rmk. For any square matrix, eigenvectors from *different eigenspaces* are *linearly independent*.

Spectral decomposition

- If a matrix A is orthogonally diagonalizable, then

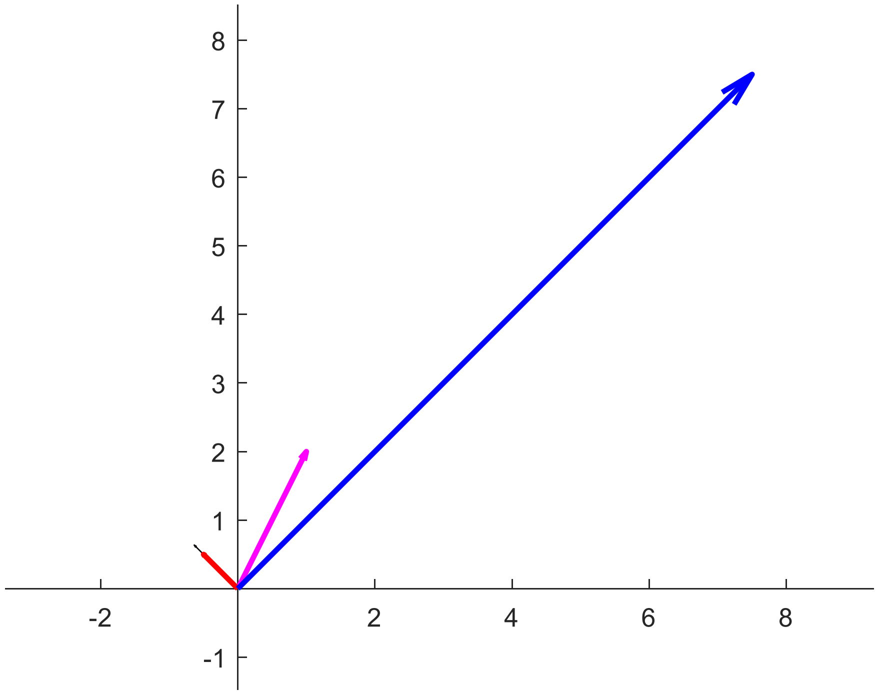
$$A = \qquad P \qquad\qquad D \qquad\qquad P^T$$

$$= \begin{bmatrix} | & & | \\ p_1 & \cdots & p_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} - & p_1^T & - \\ & \vdots & \\ - & p_n^T & - \end{bmatrix}$$

$$= \lambda_1 p_1 p_1^T + \cdots + \lambda_n p_n p_n^T$$

- which is called a spectral decomposition (or an **e**igen**v**alue **d**ecomposition).

Geometric interpretation of EVD

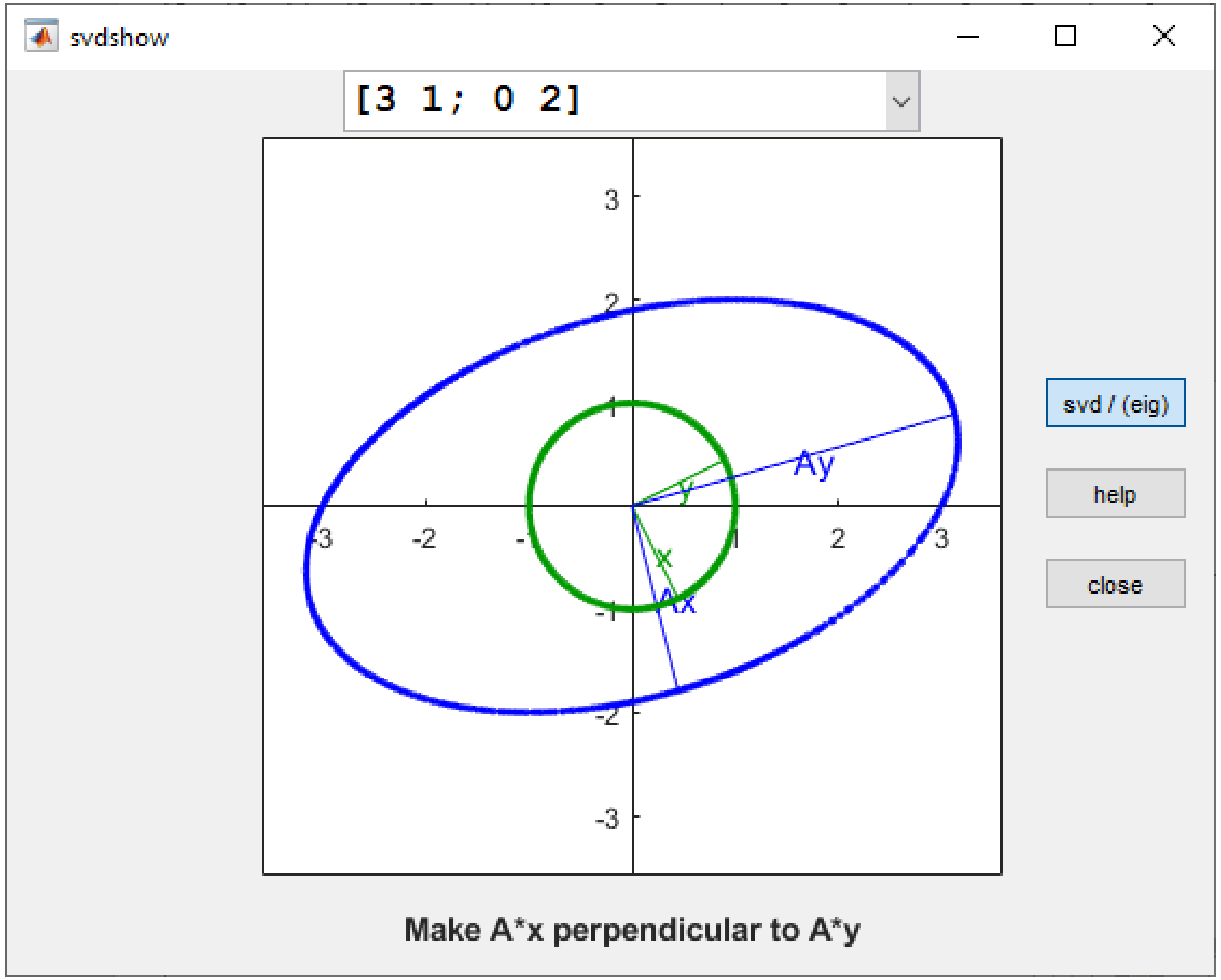


$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$$

$$= 1 \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + 5 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

- lab: example_geometric_interpretation_of_EVD.m

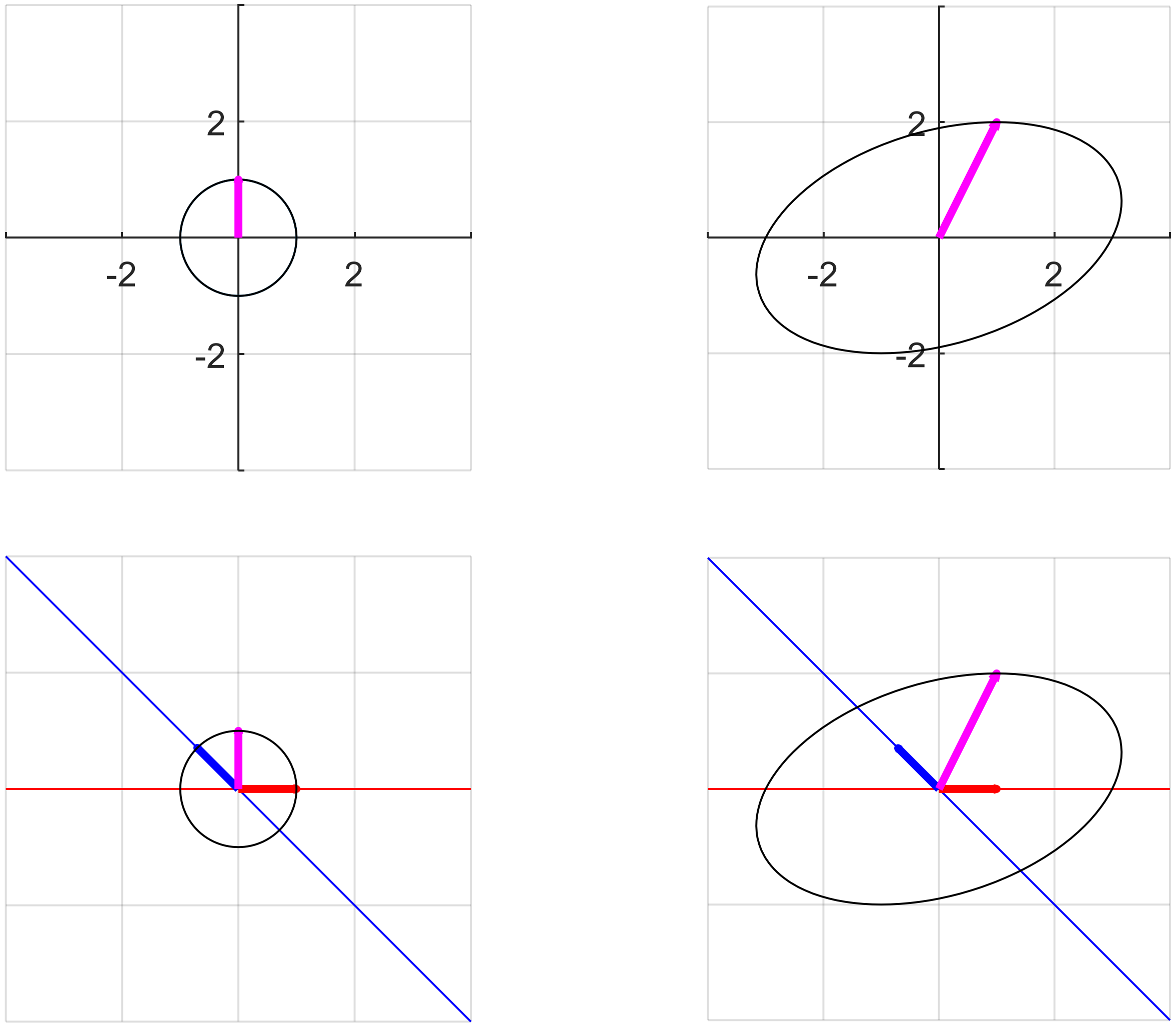
eigshow



- Q. Equation of the ellipse?
- Q. Direction of major/minor axes?

- lab: eigshow.m

Geometric interpretation of diagonalization



- lab: example_diagonalization.m

3.4 Singular Value Decomposition

Motivation of SVD

- Even if a square matrix A is not symmetric, A may be decomposed as $A = PDP^{-1}$ where P is not orthogonal.
- However, an orthogonal matrix or an orthonormal basis is too good to abandon.
- What do we have to throw away?
  - 1) $A = PHP^T$ where P is orthogonal and H is upper-triangular. (Hessenberg, Schur)
  - 2) $A = U\Sigma V^T$ where U, V are orthogonal and Σ is diagonal. (SVD)

- Can we even decompose a non-square matrix in the form of $A = U\Sigma V^T$?

Singular Value Decomposition; SVD

- [Thm] If A is an $n \times n$ matrix of rank k, then A can be factored as

$$A = U\Sigma \mathrm{V}^T$$

$$= \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$

- where U and V are $n \times n$ *orthogonal* matrices and Σ is an $n \times n$ *diagonal* matrix whose main diagonal has k positive entries and $n - k$ zeros.

Singular Value Decomposition; SVD

$$\begin{tikzcd}[column sep=huge, row sep=huge]
{[x]_S} \arrow[r, "{[T]_S = A}"] \arrow[d, "V^T"'] & {[T(x)]_S} \\
{[x]_V} \arrow[r, "{[T]_V^U = \Sigma}"'] & {[T(x)]_U} \arrow[u, "U"']
\end{tikzcd}$$

coordinate change

from S to V, or $[\mathrm{Id}]_S^V$

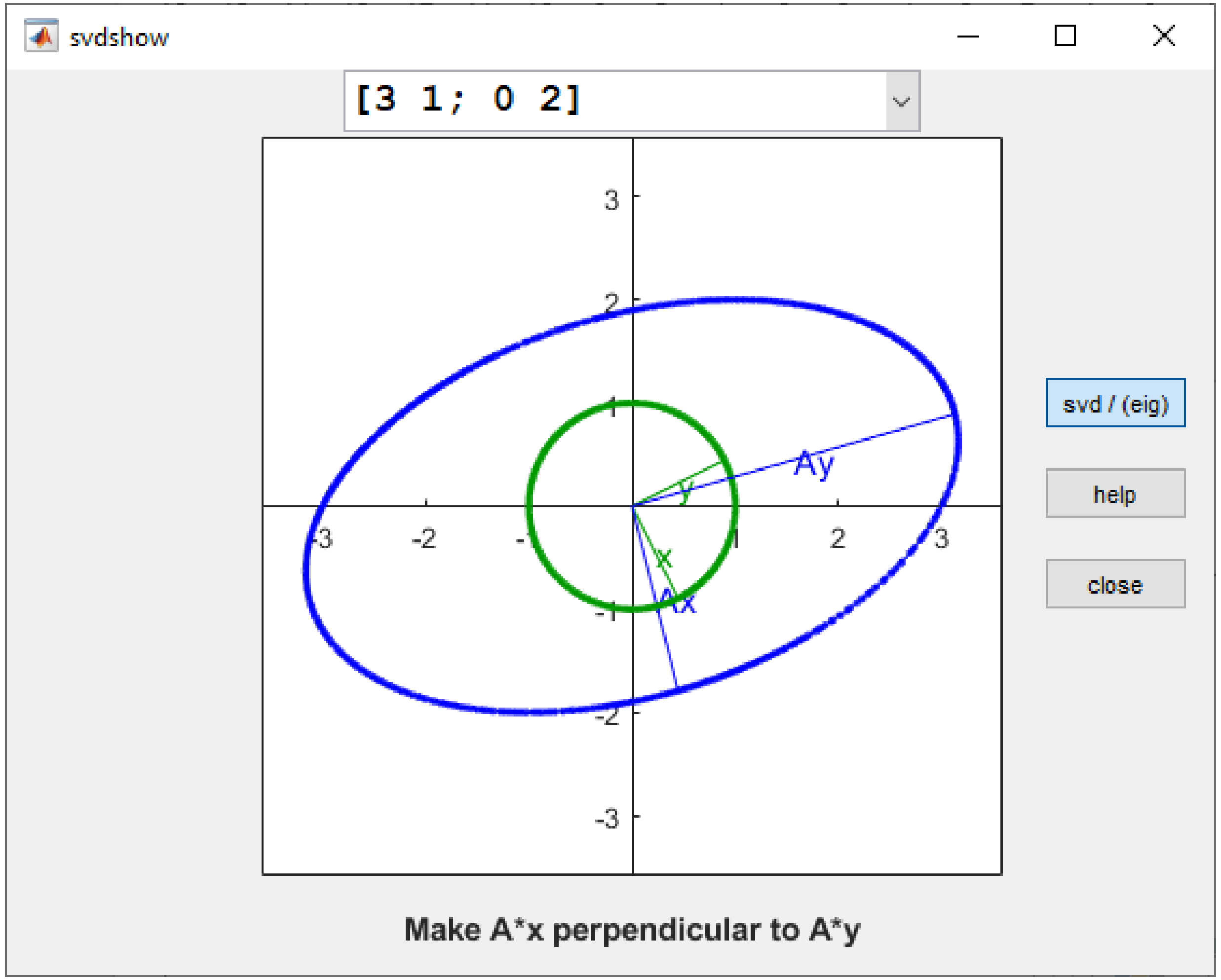
coordinate change

from U to S, or $[\mathrm{Id}]_U^S$

$$A = U\Sigma V^T$$

- Rmk. Domain and codomain of Σ have different bases.

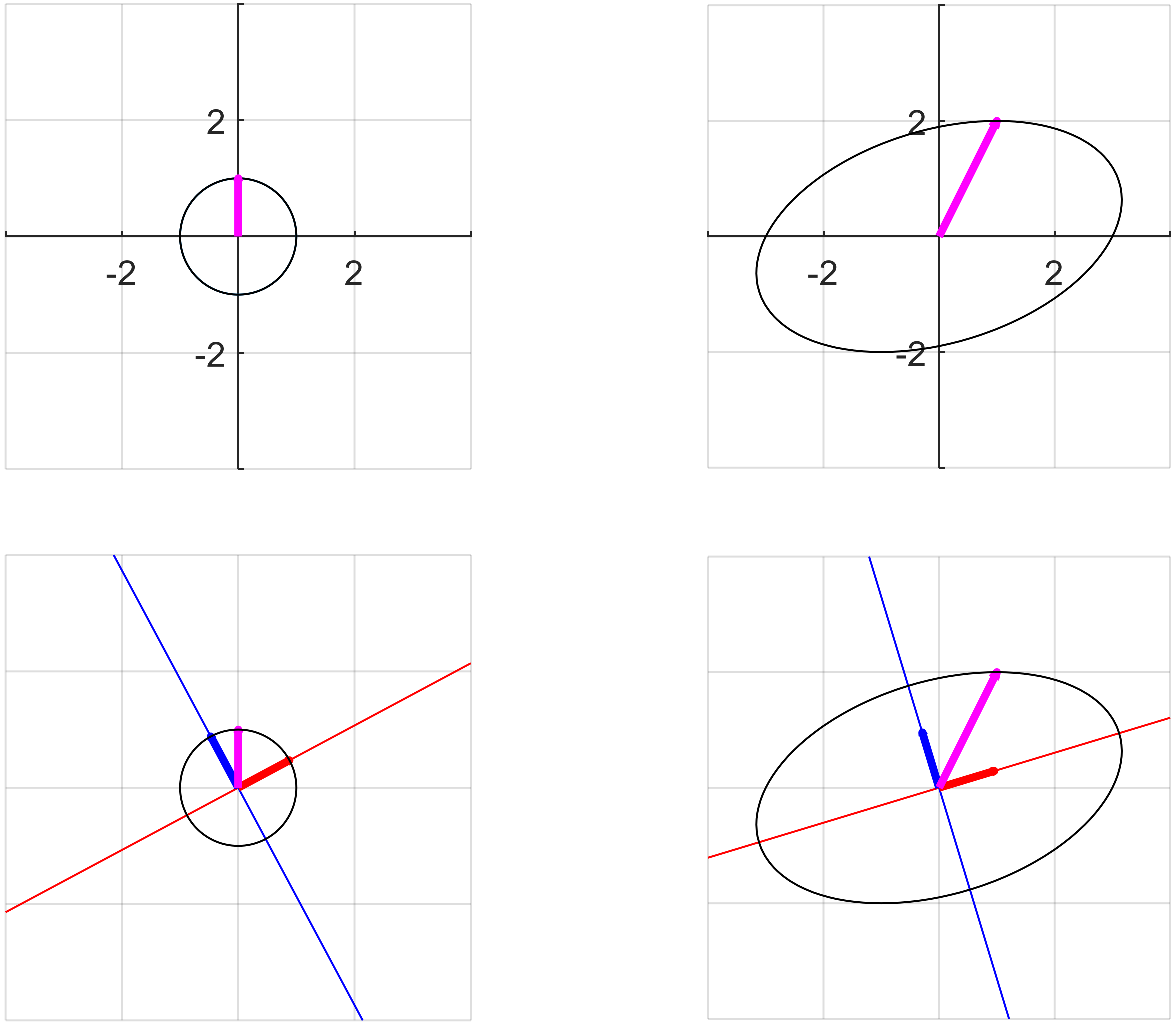
eigshow



- Q. Equation of the ellipse?
- Q. Direction of major/minor axes?

- lab: eigshow.m

Geometric interpretation of SVD



• lab: example_SVD.m

Singular Value Decomposition; SVD

$$A = U\Sigma V^T = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$

- Rmk.
  - $A^T A = VDV^T$ where D is diagonal and $\text{diag}(D) \geq 0$.
  - $\Sigma^2 = D$.
  - Main diagonal of Σ are in nonincreading order.
  - σ_i: singular values of A
  - u_i: left singular vectors of A
  - v_i: right singular vectors of A

SVD of a general matrix

- [Thm] If A is an $m \times n$ matrix of rank k, then A can be factored as

$$A = U\Sigma V^T$$

$$= \begin{bmatrix} | & & | & | & & | \\ u_1 & \cdots & u_k & u_{k+1} & \cdots & u_m \\ | & & | & | & & | \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0_{k\times(n-k)} \\ & & \sigma_k & \\ \hline & 0_{(m-k)\times k} & & 0_{(m-k)\times(n-k)} \end{array}\right] \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \\ - & v_{k+1}^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$

$$m \times m \qquad\qquad m \times n \qquad\qquad n \times n$$

Reduced SVD

$$A = U\Sigma V^T = \begin{bmatrix} | & & | & | & & | \\ u_1 & \cdots & u_k & u_{k+1} & \cdots & u_m \\ | & & | & | & & | \end{bmatrix} \begin{bmatrix} \begin{matrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{matrix} & 0_{k\times(n-k)} \\ 0_{(m-k)\times k} & 0_{(m-k)\times(n-k)} \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \\ - & v_{k+1}^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \end{bmatrix} = U_1 \Sigma_1 V_1^T$$

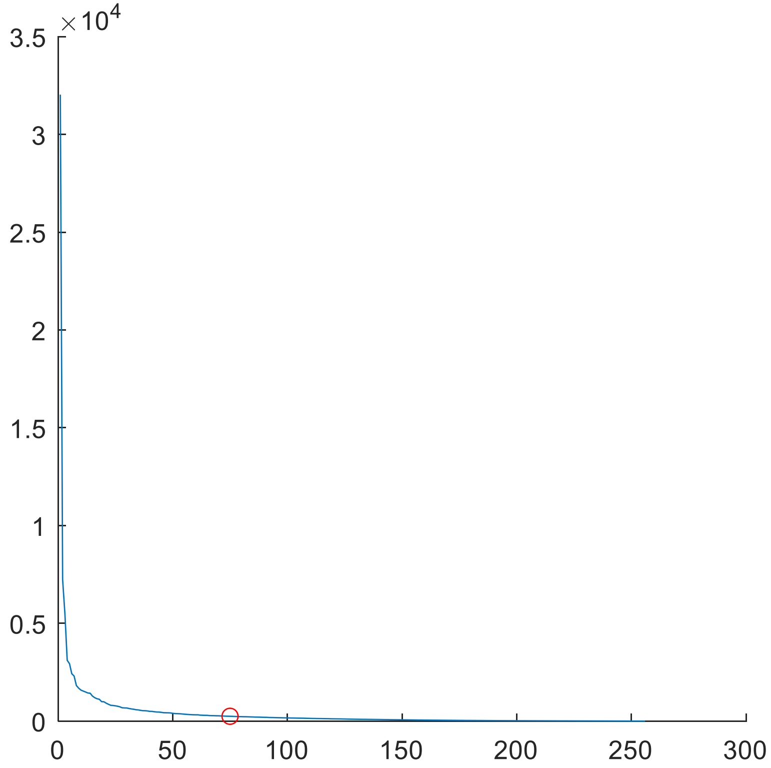
$m \times k$ $\quad$ $k \times k$ $\quad$ $k \times n$

Reduced singular value expansion

$$A = U_1\Sigma_1 V_1^T \quad = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \end{bmatrix}$$

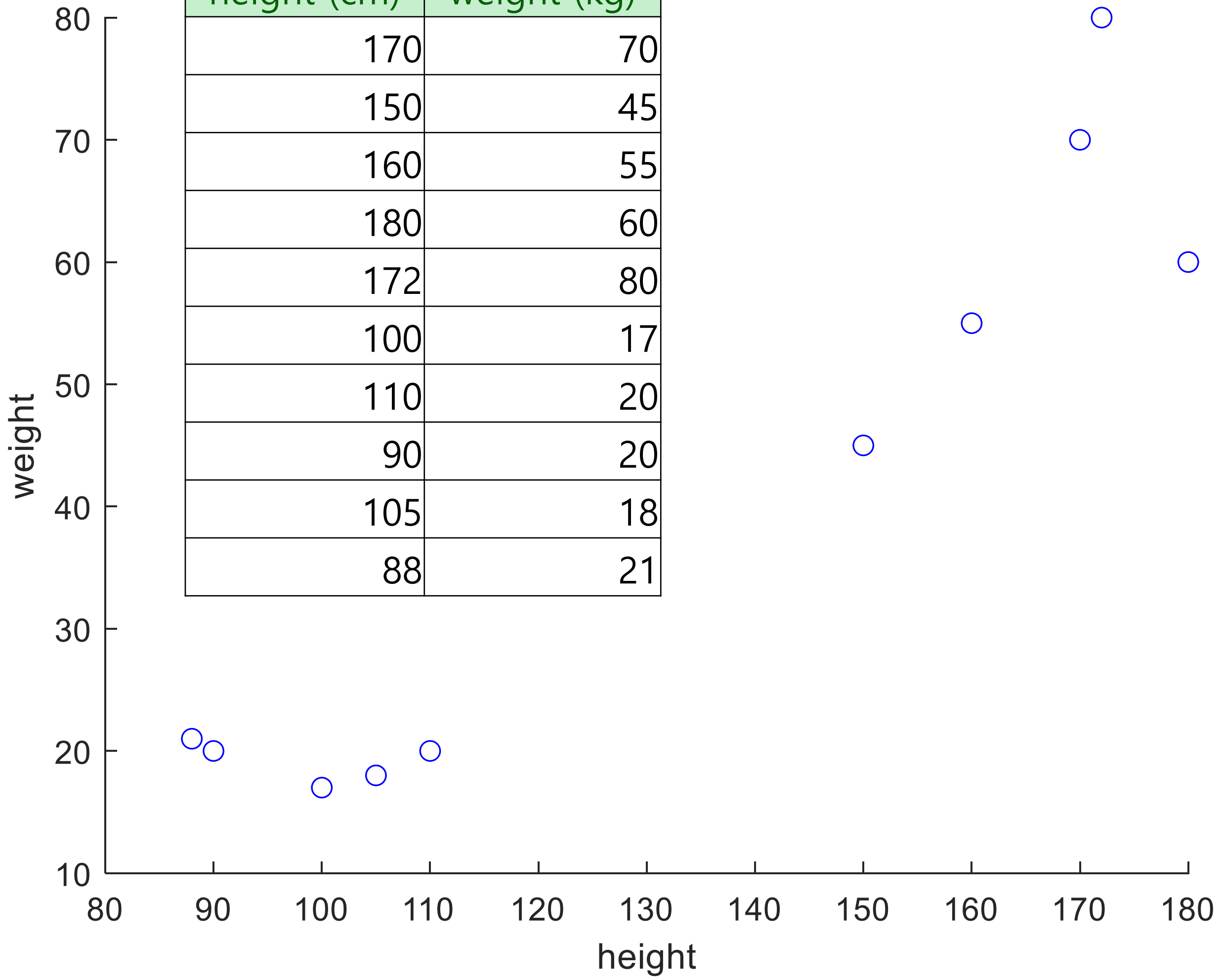
$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T$$

Image compression using SVD



- lab: example_cameraman_image_compression.m

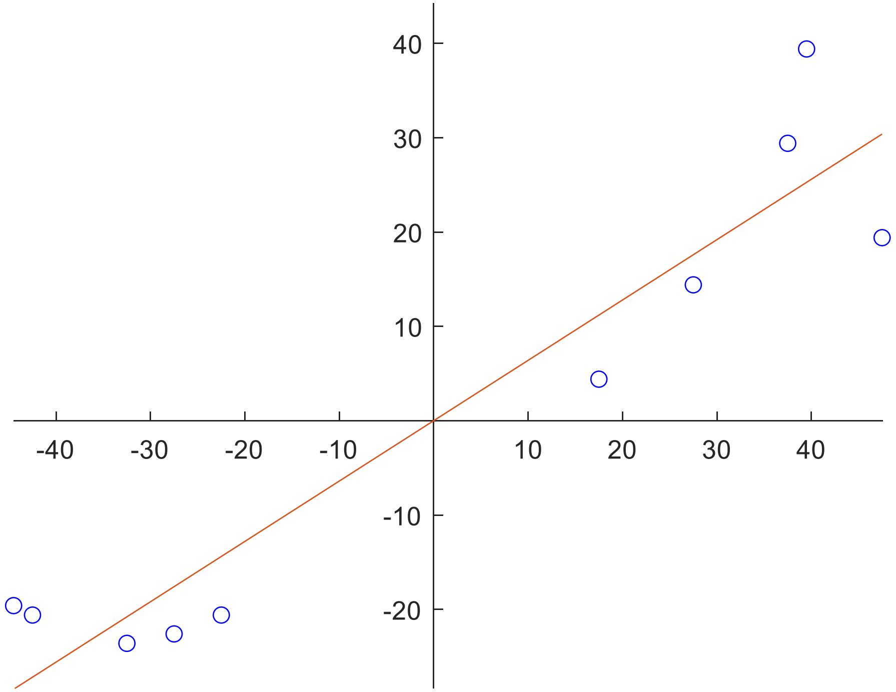
Principal Component Analysis (PCA)



height (cm)	weight (kg)
170	70
150	45
160	55
180	60
172	80
100	17
110	20
90	20
105	18
88	21

- Data for height and weight
- What is the principal axis?
- How to divide into two groups?

Principal Component Analysis (PCA)



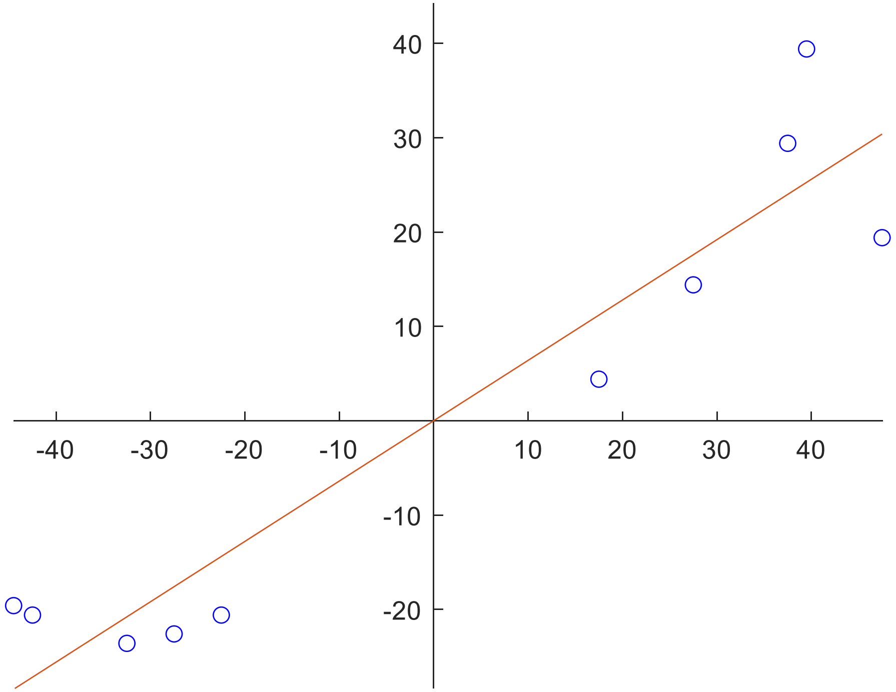
$X =$

height (cm)	weight (kg)
170	70
150	45
160	55
180	60
172	80
100	17
110	20
90	20
105	18
88	21

$$X_c = X - \mathrm{mean}(X)$$

$Y = X_c V$ ➔ Covariance matrix of Y should be diagonal.

Principal Component Analysis (PCA)



$$Y^TY = \begin{bmatrix} \mathrm{var}(y_1) & 0 \\ 0 & \mathrm{var}(y_2) \end{bmatrix}$$

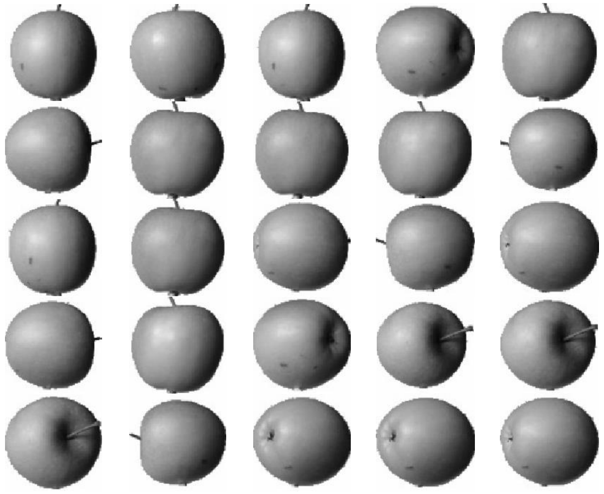
$$Y^TY = V^TX_c^TX_cV = D$$

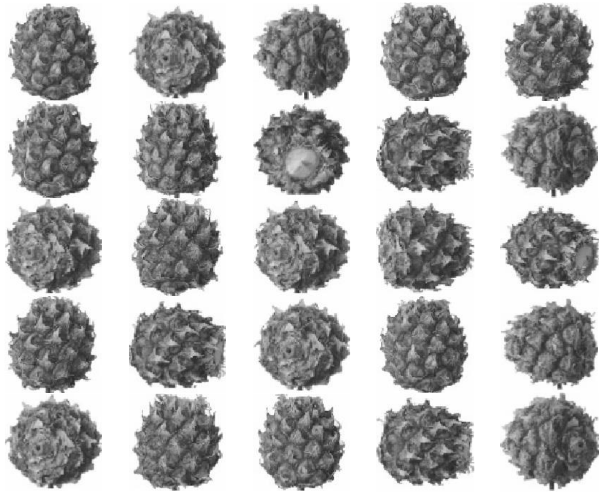
$$X_c^TX_c = VDV^T$$

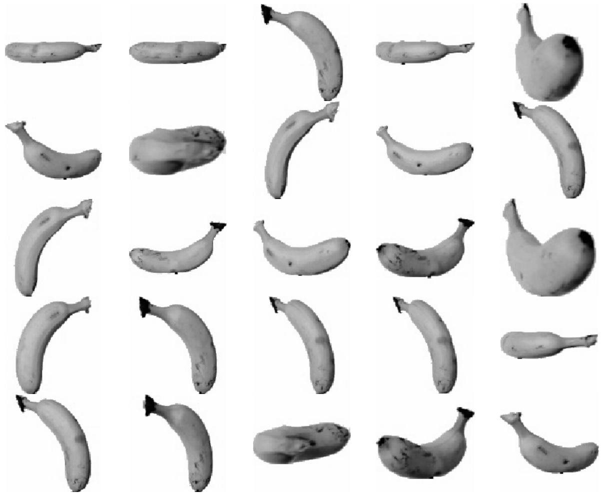
$$X_c = U\Sigma V$$

- Column vectors of V are principal axes.
- lab: example_PCA_simple.m

Principal Component Analysis (PCA)

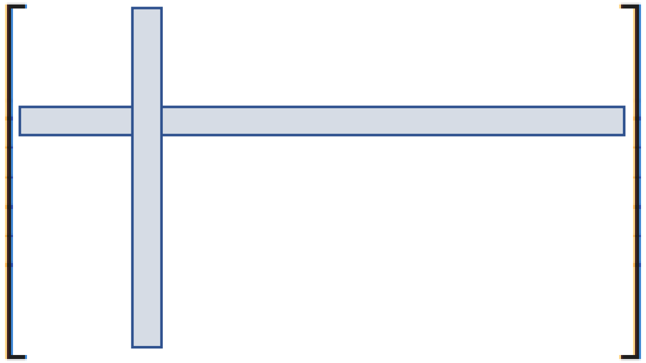






- 300 fruit images are given.
- Each image is 100 x 100.
- How to divide the fruit images into three groups?

Principal Component Analysis (PCA)

$X =$ 

flattened single image

single feature (pixel value)

$V =$ 

principal axes

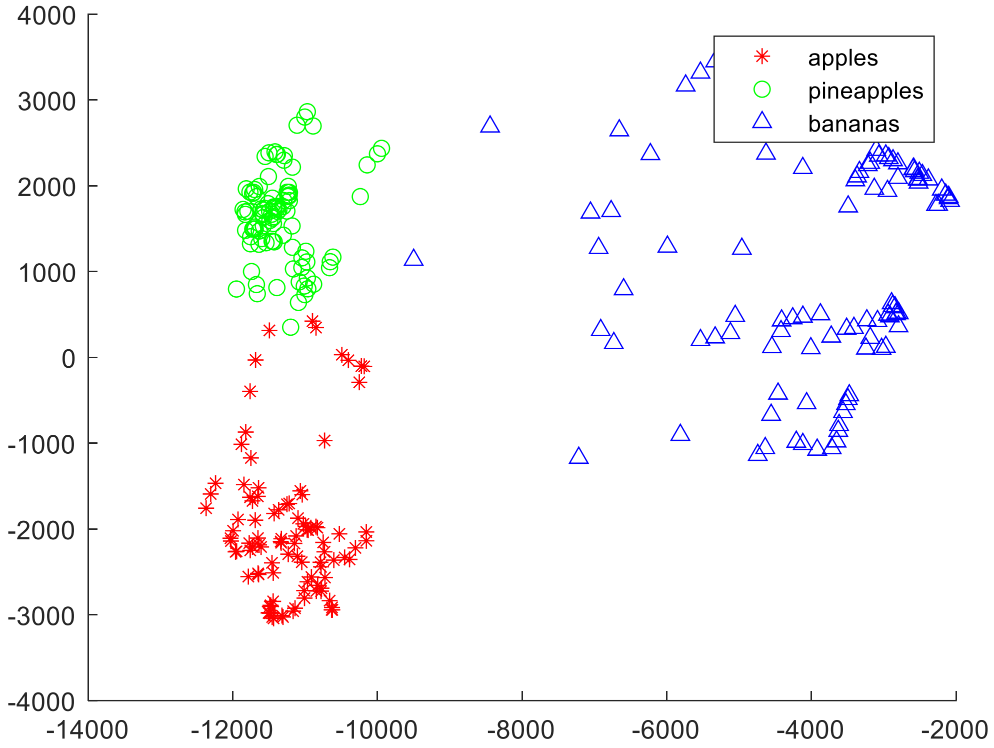
$$X_c = U\Sigma V^T$$

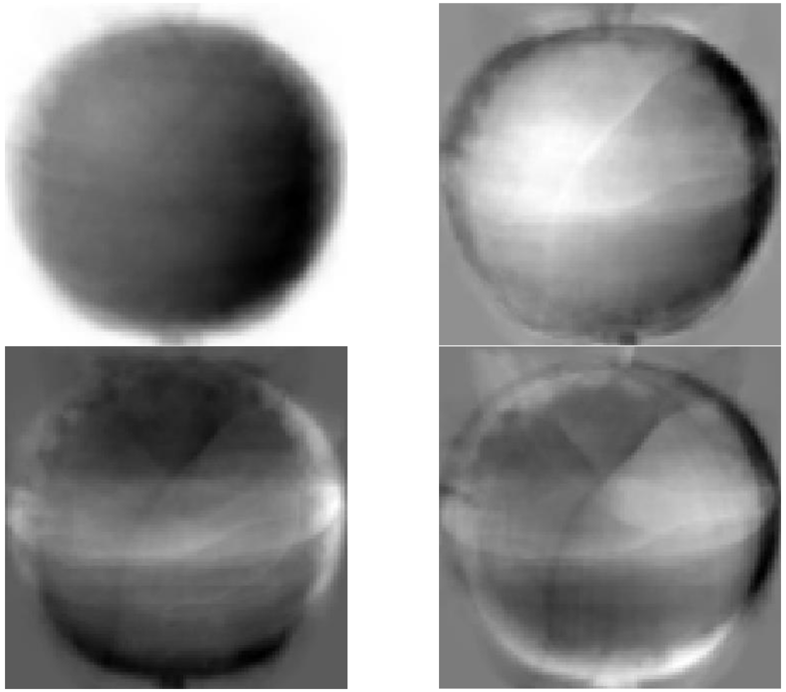
U: 300×300

Σ: 300×300

V^T: 300×10000

Principal Component Analysis (PCA)

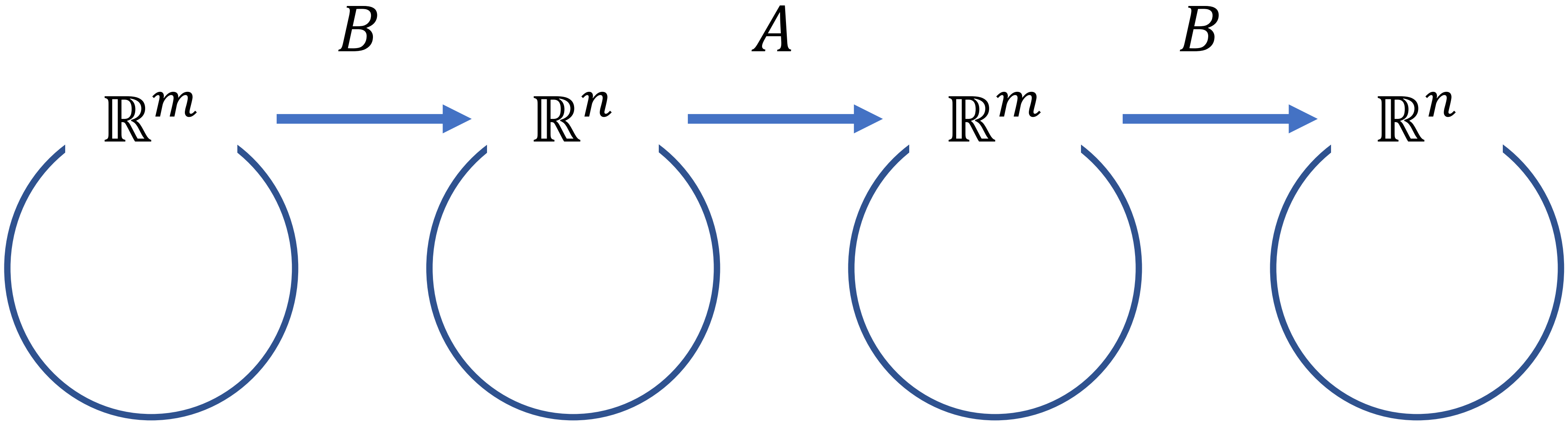




- lab: example_PCA_fruit.m

3.5 Pseudoinverse

Inverse of a *non-square* matrix



$$\mathbb{R}^m \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \xrightarrow{B} \mathbb{R}^n$$

- Let A be an $m \times n$ matrix A with full row rank $(m < n)$.
  - Is there an $n \times m$ matrix B such that $BA = I_n$?
  - Is there an $n \times m$ matrix B such that $AB = I_m$?

Inverse of a *non-square* matrix

$$
\begin{tikzcd}
{[x]_S} \arrow[rr, "{[T]_S = A}"] \arrow[d, "V^T"'] & & {[T(x)]_S} \\
{[x]_V} \arrow[rr, "{[T]_V^U = \Sigma}"'] & & {[T(x)]_U} \arrow[u, "U"']
\end{tikzcd}
$$

$$A = \underset{m \times m}{U}\,\underset{m \times n}{\Sigma}\,\underset{n \times n}{V^T}$$

- Reversing the direction does not work because Σ is not square.

$$
\begin{tikzcd}
{[x]_S} \arrow[rr, "{[T]_S = A}"] \arrow[d, "V_1^T"'] & & {[T(x)]_S} \\
{[x]_V} \arrow[rr, "{[T]_V^U = \Sigma_1}"'] & & {[T(x)]_U} \arrow[u, "U_1"']
\end{tikzcd}
$$

$$A = \underset{m \times n}{U_1}\,\underset{n \times n}{\Sigma_1}\,\underset{n \times n}{V_1^T}$$

- Reversing the direction *works* for the *reduced* SVD because A has full row rank.

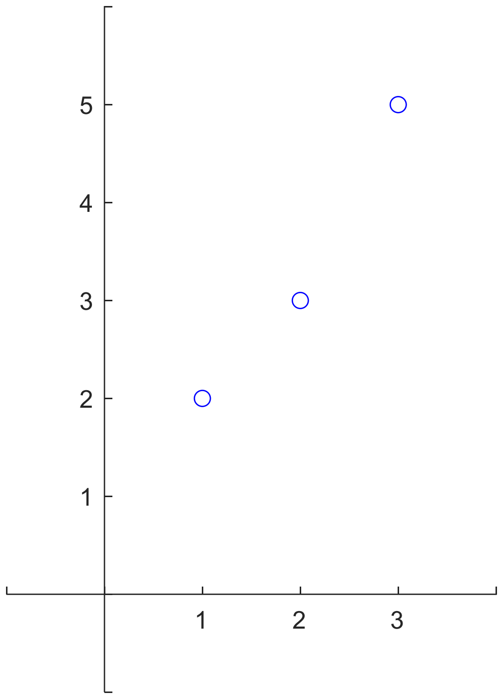
Inverse of a *non-square* matrix

- Let A be an $m \times n$ matrix, $A = U_1 \Sigma_1 V_1^T$ be the SVD of A, $A^+ = V_1 \Sigma_1^{-1} U_1^T$.
- If A is full row rank, $AA^+ = I_m$, which is called the right inverse of A.
- If A is full column rank, $A^+ A = I_n$, which is called the left inverse of A.
- [Thm] If A is an $m \times n$ matrix with full column rank, then
$$A^+ = (A^T A)^{-1} A^T$$
- If A has full row rank, then
$$A^+ = A^T (AA^T)^{-1}$$

Properties of pseudoinverse

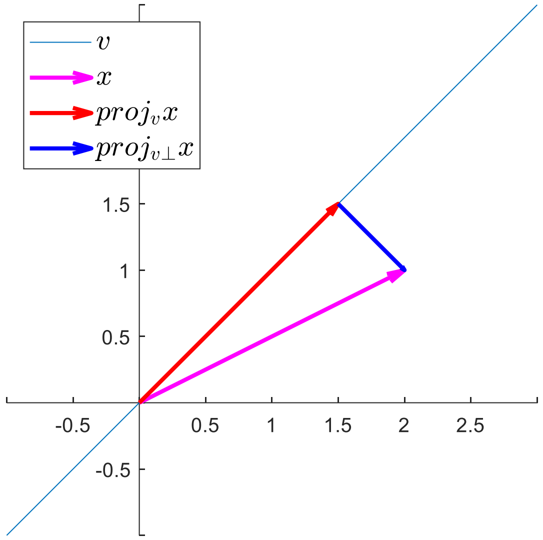
- If A^+ is the pseudoinverse of an $m \times n$ matrix A, then:
  - (a) $AA^+A = A$
  - (b) $A^+AA^+ = A^+$
  - (c) $(AA^+)^T = AA^+$
  - (d) $(A^+A)^T = A^+A$
  - (e) $(A^T)^+ = (A^+)^T$
  - (f) $A^{++} = A$

Pseudoinverse solves the Least Square Prob.



$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$A \quad x \quad + \quad e \quad = \quad y$$



$$\text{proj}_{\text{col}(A)} y = A(A^T A)^{-1} A^T y$$

$$\hat{y} = Ax$$

$$x = (A^T A)^{-1} A^T y = A^+ y$$

- AA^+x is the projection of x onto $\text{col}(A)$.
- Best fit is projecting y onto $\text{col}(A)$.

- lab: example_least_square_is_projection.m

Condition number

$$
\begin{tikzcd}
{[x]_S} \arrow[r, "{[T]_S = A}"] \arrow[d, "V^T"'] & {[T(x)]_S} \\
{[x]_V} \arrow[r, "{[T]_V^U = \Sigma}"'] & {[T(x)]_U} \arrow[u, "U"']
\end{tikzcd}
$$

- What matters is Σ which stretches/compresses a vector in each direction.
- [Def] **Condition number** of a matrix A is the ratio of the largest singular value of A to the smallest singular value of A.

$$\mathrm{cond}(A) = \frac{\sigma_1}{\sigma_k}$$

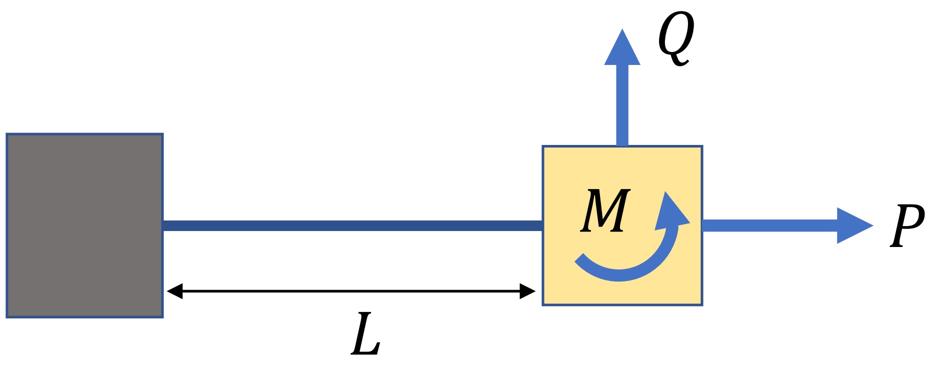
- Condition number of A is too large.
  ≡ Some dimension of $\mathrm{Im}(A)$ is negligible.
  ≡ A is near-singular.

- lab: eigshow.m

Multibody vibration problem

is an eigenvalue problem.

Compliant mechanisms



$$\delta_x = \frac{PL}{EA}$$

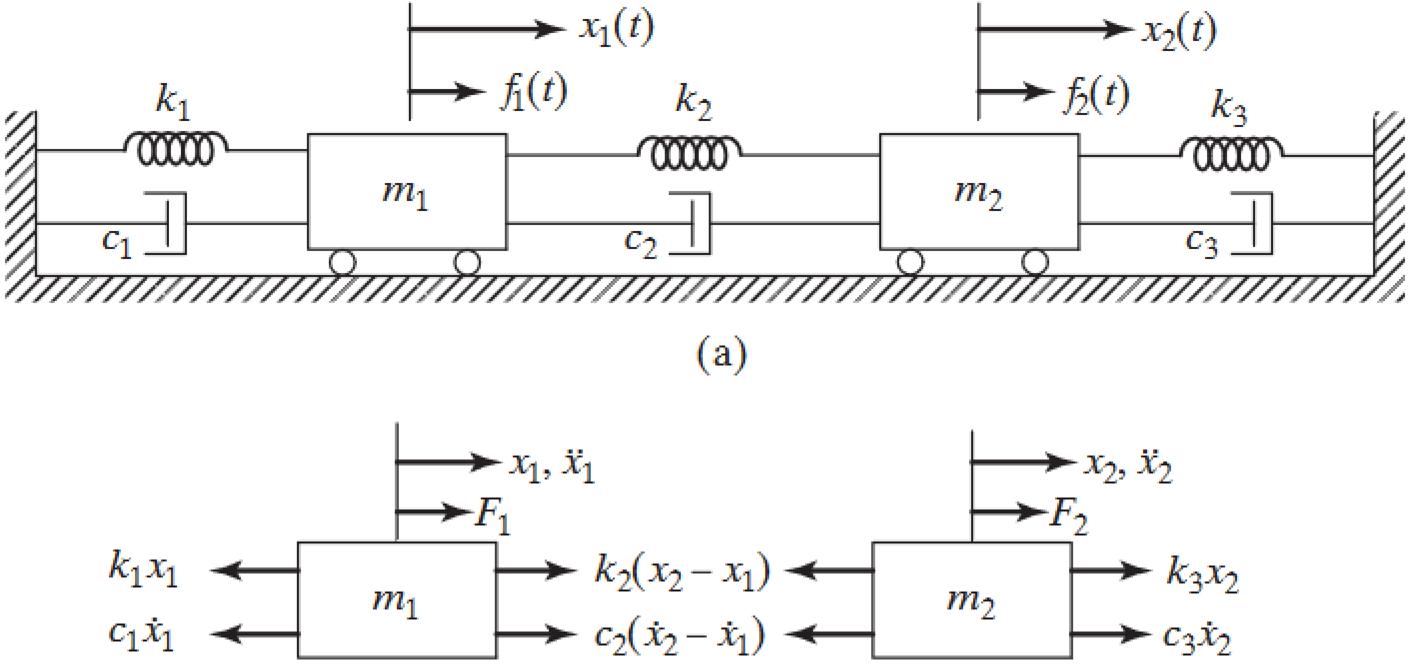
$$\delta_y = \frac{QL^3}{3EI} + \frac{ML^2}{2EI}$$

$$\theta_z = \frac{QL^2}{2EI} + \frac{ML}{EI}$$

$$\begin{bmatrix} \delta_x \\ \delta_y \\ \theta_z \end{bmatrix} = \begin{bmatrix} \frac{L}{EA} & 0 & 0 \\ 0 & \frac{L^3}{3EI} & \frac{L^2}{2EI} \\ 0 & \frac{L^2}{2EI} & \frac{L}{EI} \end{bmatrix} \begin{bmatrix} P \\ Q \\ M \end{bmatrix}$$

$$x = Cf \qquad f = Kx$$

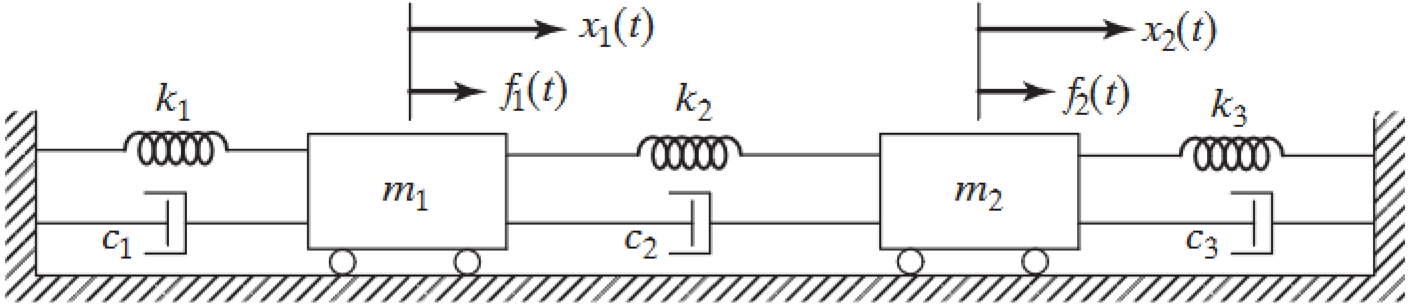
2-DOF system



$$m_1\ddot{x}_1(t) + (k_1 + k_2)x_1(t) - k_2x_2(t) = 0 \qquad x_1(t) = X_1\cos(\omega t + \phi)$$

$$m_2\ddot{x}_2(t) - k_2x_1(t) + (k_2 + k_3)x_2(t) = 0 \qquad x_2(t) = X_2\cos(\omega t + \phi)$$

2-DOF system



$$M\ddot{x} + Kx = 0$$

$$(K - \omega^2 M)x = 0$$

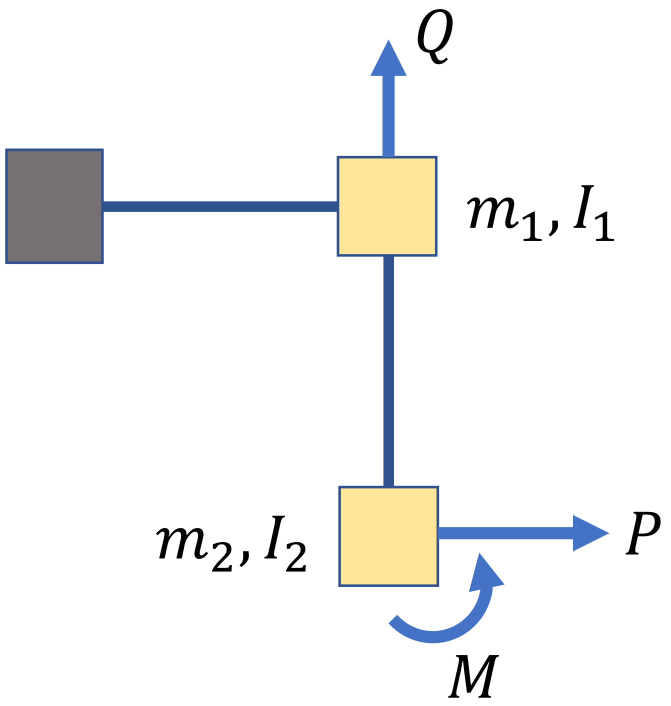
$$KM^{-1}x = \omega^2 x$$

```
[V, D] = eig(K*inv(M));

[V, D] = eig(inv(sqrt(M))*K*inv(sqrt(M)));

[V, D] = eig(K, M);
```

Multi-mass Compliant Mechanism



$$M\ddot{x} + Kx = 0$$

```
[V, D] = eig(K*inv(M));

[V, D] = eig(inv(sqrt(M))*K*inv(sqrt(M)));

[V, D] = eig(K, M);
```

See also.
https://ocw.snu.ac.kr/node/30658