# Chapter 05
# Linear Least-squares

Lin ZHANG
School of Computer Science and Technology
Tongji University

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Matrix differentiation

- Function is a vector and the variable is a scalar

$$\boldsymbol{f}(t) = \left[ f_1(t), f_2(t), \ldots, f_n(t) \right]^T$$

Definition

$$\frac{d\boldsymbol{f}}{dt} = \left[ \frac{df_1(t)}{dt}, \frac{df_2(t)}{dt}, \ldots, \frac{df_n(t)}{dt} \right]^T$$

# Matrix differentiation

- Function is a matrix and the variable is a scalar

$$\boldsymbol{F}(t)=\begin{bmatrix} f_{11}(t)\ f_{12}(t),\dots,f_{1m}(t) \\ f_{21}(t)\ f_{22}(t),\dots,f_{2m}(t) \\ \vdots \\ f_{n1}(t)\ f_{n2}(t),\dots,f_{nm}(t) \end{bmatrix}=\left[f_{ij}(t)\right]_{n\times m}$$

Definition

$$\frac{d\boldsymbol{F}}{dt}=\begin{bmatrix} \frac{df_{11}(t)}{dt}\ \frac{df_{12}(t)}{dt},\dots,\frac{df_{1m}(t)}{dt} \\ \frac{df_{21}(t)}{dt}\ \frac{df_{22}(t)}{dt},\dots,\frac{df_{2m}(t)}{dt} \\ \vdots \\ \frac{df_{n1}(t)}{dt}\ \frac{df_{n2}(t)}{dt},\dots,\frac{df_{nm}(t)}{dt} \end{bmatrix}=\left[\frac{df_{ij}(t)}{dt}\right]_{n\times m}$$

# Matrix differentiation

- Function is a scalar and the variable is a vector

$$f(\boldsymbol{x}), \boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$$

Definition

$$\frac{df}{d\boldsymbol{x}} = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T$$

In a similar way,

$$f(\boldsymbol{x}), \boldsymbol{x} = (x_1, x_2, \ldots, x_n)$$

$$\frac{df}{d\boldsymbol{x}} = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]$$

# Matrix differentiation

- Function is a vector and the variable is a vector

$$\boldsymbol{x}=\left[x_1,x_2,\ldots,x_n\right]^T, \boldsymbol{y}=\left[y_1(\boldsymbol{x}),y_2(\boldsymbol{x}),\ldots,y_m(\boldsymbol{x})\right]^T$$

Definition

$$\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}=\begin{bmatrix} \frac{\partial y_1(\boldsymbol{x})}{\partial x_1},\frac{\partial y_1(\boldsymbol{x})}{\partial x_2},\ldots,\frac{\partial y_1(\boldsymbol{x})}{\partial x_n} \\ \frac{\partial y_2(\boldsymbol{x})}{\partial x_1},\frac{\partial y_2(\boldsymbol{x})}{\partial x_2},\ldots,\frac{\partial y_2(\boldsymbol{x})}{\partial x_n} \\ \vdots \\ \frac{\partial y_m(\boldsymbol{x})}{\partial x_1},\frac{\partial y_m(\boldsymbol{x})}{\partial x_2},\ldots,\frac{\partial y_m(\boldsymbol{x})}{\partial x_n} \end{bmatrix}_{m\times n}$$

# Matrix differentiation

- Function is a vector and the variable is a vector

$$\boldsymbol{x}=\left[x_1,x_2,\ldots,x_n\right]^T,\ \boldsymbol{y}=\left[y_1(\boldsymbol{x}),y_2(\boldsymbol{x}),\ldots,y_m(\boldsymbol{x})\right]^T$$

In a similar way,

$$\frac{d\boldsymbol{y}^T}{d\boldsymbol{x}}=\begin{bmatrix}\frac{\partial y_1(\boldsymbol{x})}{\partial x_1},\frac{\partial y_2(\boldsymbol{x})}{\partial x_1},\ldots,\frac{\partial y_m(\boldsymbol{x})}{\partial x_1}\\ \frac{\partial y_1(\boldsymbol{x})}{\partial x_2},\frac{\partial y_2(\boldsymbol{x})}{\partial x_2},\ldots,\frac{\partial y_m(\boldsymbol{x})}{\partial x_2}\\ \vdots\\ \frac{\partial y_1(\boldsymbol{x})}{\partial x_n},\frac{\partial y_2(\boldsymbol{x})}{\partial x_n},\ldots,\frac{\partial y_m(\boldsymbol{x})}{\partial x_n}\end{bmatrix}_{n\times m}$$

# Matrix differentiation

- Function is a vector and the variable is a vector

Example:

$$\boldsymbol{y}=\begin{bmatrix} y_1(\boldsymbol{x}) \\ y_2(\boldsymbol{x}) \end{bmatrix}, \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, y_1(\boldsymbol{x})=x_1^2-x_2, y_2(\boldsymbol{x})=x_3^2+3x_2$$

$$\frac{d\boldsymbol{y}^T}{d\boldsymbol{x}}=\begin{bmatrix} \frac{\partial y_1(\boldsymbol{x})}{\partial x_1} & \frac{\partial y_2(\boldsymbol{x})}{\partial x_1} \\ \frac{\partial y_1(\boldsymbol{x})}{\partial x_2} & \frac{\partial y_2(\boldsymbol{x})}{\partial x_2} \\ \frac{\partial y_1(\boldsymbol{x})}{\partial x_3} & \frac{\partial y_2(\boldsymbol{x})}{\partial x_3} \end{bmatrix}=\begin{bmatrix} 2x_1 & 0 \\ -1 & 3 \\ 0 & 2x_3 \end{bmatrix}$$

# Matrix differentiation

- Function is a scalar and the variable is a matrix

$$f(\boldsymbol{X}), \boldsymbol{X} \in \mathbb{R}^{m \times n}$$

Definition

$$\frac{df(\boldsymbol{X})}{d\boldsymbol{X}} = \begin{bmatrix} \frac{\partial f}{\partial x_{11}} \frac{\partial f}{\partial x_{12}} \cdots \frac{\partial f}{\partial x_{1n}} \\ \cdots \\ \frac{\partial f}{\partial x_{m1}} \frac{\partial f}{\partial x_{m2}} \cdots \frac{\partial f}{\partial x_{mn}} \end{bmatrix}$$

# Matrix differentiation

- Useful results

(1) $\boldsymbol{x}, \boldsymbol{a} \in \mathbb{R}^{n\times 1}$

Then,

$$\frac{d\boldsymbol{a}^T\boldsymbol{x}}{d\boldsymbol{x}} = \boldsymbol{a}, \frac{d\boldsymbol{x}^T\boldsymbol{a}}{d\boldsymbol{x}} = \boldsymbol{a}$$

*How to prove?*

# Matrix differentiation

- Useful results

(2) $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$ Then, $\dfrac{d\boldsymbol{x}^T\boldsymbol{x}}{d\boldsymbol{x}} = 2\boldsymbol{x}$

(3) $\boldsymbol{y}(\boldsymbol{x}) \in \mathbb{R}^{m\times 1}$, $\boldsymbol{x} \in \mathbb{R}^{n\times 1}$, $\dfrac{d\boldsymbol{y}^T(\boldsymbol{x})}{d\boldsymbol{x}} = \left(\dfrac{d\boldsymbol{y}(\boldsymbol{x})}{d\boldsymbol{x}^T}\right)^T$

(4) $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}$ Then, $\dfrac{d\boldsymbol{A}\boldsymbol{x}}{d\boldsymbol{x}^T} = \boldsymbol{A}$

(5) $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}$ Then, $\dfrac{d\boldsymbol{x}^T\boldsymbol{A}^T}{d\boldsymbol{x}} = \boldsymbol{A}^T$

(6) $\boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}$ Then, $\dfrac{d\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{d\boldsymbol{x}} = (\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x}$

(7) $\boldsymbol{X} \in \mathbb{R}^{m\times n}, \boldsymbol{a} \in \mathbb{R}^{m\times 1}, \boldsymbol{b} \in \mathbb{R}^{n\times 1}$ Then, $\dfrac{d\boldsymbol{a}^T\boldsymbol{X}\boldsymbol{b}}{d\boldsymbol{X}} = \boldsymbol{a}\boldsymbol{b}^T$

# Matrix differentiation

- Useful results

(8) $\boldsymbol{X} \in \mathbb{R}^{n\times m}, \boldsymbol{a} \in \mathbb{R}^{m\times 1}, \boldsymbol{b} \in \mathbb{R}^{n\times 1}$ Then, $\frac{d\boldsymbol{a}^T\boldsymbol{X}^T\boldsymbol{b}}{d\boldsymbol{X}} = \boldsymbol{b}\boldsymbol{a}^T$

(9) $\boldsymbol{X} \in \mathbb{R}^{m\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times m}$ Then, $\frac{d\left(tr\left(\boldsymbol{X}\boldsymbol{B}\right)\right)}{d\boldsymbol{X}} = \boldsymbol{B}^T$

(10) $\boldsymbol{X} \in \mathbb{R}^{n\times n}, \boldsymbol{X}$ is invertible, $\frac{d\left|\boldsymbol{X}\right|}{d\boldsymbol{X}} = \left|\boldsymbol{X}\right|\left(\boldsymbol{X}^{-1}\right)^T$

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Lagrange multiplier

- Single-variable function

$f(x)$ is differentiable in $(a, b)$. At $x_0 \in (a,b)$, $f(x)$ achieves an extremum

$$\Rightarrow \frac{df}{dx}\Big|_{x_0} = 0$$

- Two-variables function

$f(x, y)$ is differentiable in its domain. At $(x_0, y_0)$, $f(x, y)$ achieves an extremum

$$\Rightarrow \frac{\partial f}{\partial x}\Big|_{(x_0, y_0)} = 0, \frac{\partial f}{\partial y}\Big|_{(x_0, y_0)} = 0$$

# Lagrange multiplier

- In general case

If $f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^{n\times 1}$ achieves a local extremum at $\boldsymbol{x}_0$ and it is derivable at $\boldsymbol{x}_0$, then $\boldsymbol{x}_0$ is a stationary point of $f(\boldsymbol{x})$, i.g.,

$$\frac{\partial f}{\partial x_1}|_{\boldsymbol{x}_0} = 0, \frac{\partial f}{\partial x_2}|_{\boldsymbol{x}_0} = 0, \ldots, \frac{\partial f}{\partial x_n}|_{\boldsymbol{x}_0} = 0$$

Or in other words,

$$\nabla f(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_0} = \boldsymbol{0}$$

# Lagrange multiplier

- Lagrange multiplier is a strategy for finding **all the possible** extremum points of a function subject to equality constraints

Problem: find **all the possible** extremum points for $y = f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^{n\times 1}$

under $m$ constraints $g_k(\boldsymbol{x}) = 0, k = 1, 2, \ldots, m$

Solution: $F(\boldsymbol{x}; \lambda_1, \ldots, \lambda_m) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x})$

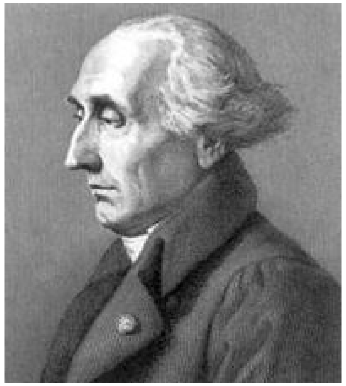
If $\boldsymbol{x}_0$ is an extremum point of $f(\boldsymbol{x})$ under constraints

$\Downarrow$

$\exists \lambda_{10}, \lambda_{20} \ldots, \lambda_{m0}$, making $(\boldsymbol{x}_0, \lambda_{10}, \lambda_{20} \ldots, \lambda_{m0})$

a stationary point of $F$

Thus, by identifying the stationary points of $F$, we can get all the possible extremum points of $f(\boldsymbol{x})$ under equality constraints



Joseph-Louis Lagrange
Jan. 25, 1736~Apr.10, 1813

# Lagrange multiplier

- Lagrange multiplier is a strategy for finding **all the possible** extremum points of a function subject to equality constraints

Problem: find **all the possible** extremum points for $y = f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^{n\times 1}$

under $m$ constraints $g_k(\boldsymbol{x}) = 0, k = 1, 2, \ldots, m$

Solution: $F(\boldsymbol{x}; \lambda_1, \ldots, \lambda_m) = f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k g_k(\boldsymbol{x})$

$(\boldsymbol{x}_0, \lambda_{10}, \ldots, \lambda_{m0})$ is a stationary point of $F$ $\Rightarrow$

$$\underbrace{\frac{\partial F}{\partial x_1} = 0, \frac{\partial F}{\partial x_2} = 0, \ldots, \frac{\partial F}{\partial x_n} = 0, \frac{\partial F}{\partial \lambda_1} = 0, \frac{\partial F}{\partial \lambda_2} = 0, \ldots, \frac{\partial F}{\partial \lambda_m} = 0}_{n+m \text{ equations!}}$$

at that point

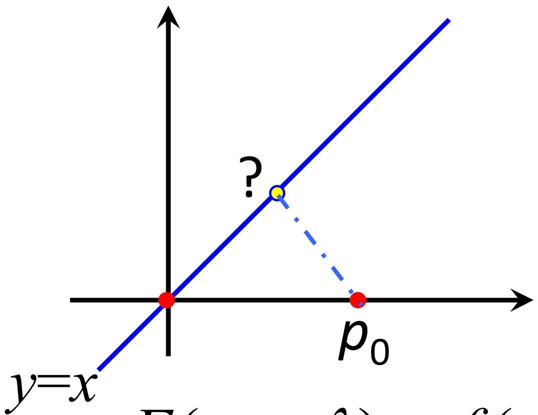
$\Downarrow$

$\boldsymbol{x}_0$ is a possible extremum point of $f(\boldsymbol{x})$ under equality constraints

# Lagrange multiplier

- Example

Problem: for a given point $p_0$ = (1, 0), among all the points lying on the line $y=x$, identify the one having the least distance to $p_0$.




The distance is

$$f(x,y)=(x-1)^2+(y-0)^2$$

Now we want to find the global minimizer of $f(x, y)$ under the constraint

$$g(x,y)=y-x=0$$

According to Lagrange multiplier method, construct the Lagrange function

$$F(x,y,\lambda)=f(x,y)+\lambda g(x,y)=(x-1)^2+y^2+\lambda(y-x)$$

Find the stationary point of $F(x,y,\lambda)$

# Lagrange multiplier

- Example

$$\begin{cases} \dfrac{\partial F}{\partial x} = 0 \\ \dfrac{\partial F}{\partial y} = 0 \\ \dfrac{\partial F}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} 2(x-1) + \lambda = 0 \\ 2y - \lambda = 0 \\ x - y = 0 \end{cases} \Rightarrow \begin{cases} x = 0.5 \\ y = 0.5 \\ \lambda = 1 \end{cases}$$

Thus, $(0.5, 0.5, 1)$ is the only stationary point of $F(x, y, \lambda)$

$(0.5,0.5)$ is the only possible extremum point of $f(x,y)$ under constraints

The global minimizer of $f(x,y)$ under constraints exists

$(0.5,0.5)$ is the global minimizer of $f(x,y)$ under constraints

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Singular Value Decomposition (SVD)

SVD decomposition theorem: Any matrix $\boldsymbol{A}_{m\times n}$ can be decomposed as the following form,

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

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are two orthogonal matrices, $rank(\boldsymbol{A})=r$,

$$\Sigma_{m\times n}=\begin{bmatrix}\Sigma_r & \boldsymbol{O}_{r\times(n-r)}\\ \boldsymbol{O}_{(m-r)\times r} & \boldsymbol{O}_{(m-r)\times(n-r)}\end{bmatrix}_{m\times n}=\begin{bmatrix}\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{bmatrix} & \boldsymbol{O}_{r\times(n-r)}\\ \boldsymbol{O}_{(m-r)\times r} & \boldsymbol{O}_{(m-r)\times(n-r)}\end{bmatrix}_{m\times n}$$

$\sigma_1,\sigma_2,\ldots,\sigma_r>0$ are called the **singular values** of $\boldsymbol{A}$

In general case, $\Sigma_{m\times n}$ is not unique. However, if $\{\sigma_i\}_{i=1}^{r}$ are arranged in order, $\Sigma_{m\times n}$ is uniquely determined by $\boldsymbol{A}$. In the following, we require that $\sigma_1\geq\sigma_2\geq,\ldots,\geq\sigma_r>0$

## Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Homography estimation—Modeled as a homogeneous linear LS problem

Suppose that by matching keypoints on $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$, we get $\mathcal{S}=\left\{\boldsymbol{x}_i \leftrightarrow \boldsymbol{x}_i^{'}\right\}_{i=1}^{p}$

where $\boldsymbol{x}_i \leftrightarrow \boldsymbol{x}_i^{'}$ means that the point $\boldsymbol{x}_i$ from $\boldsymbol{I}_1$ and the point $\boldsymbol{x}_i^{'}$ from $\boldsymbol{I}_2$ matches, and $p$ is the number of correspondence pairs

*temporarily suppose there is no outliers in* $\mathcal{S}$

All the pairs can be linked via a universal 2D linear geometric transformation $\boldsymbol{H}_{3\times 3}$

*If there is no prior knowledge about the geometric transformation between* $\boldsymbol{I}_1$ *and* $\boldsymbol{I}_2$

$\boldsymbol{H}$ can take the most general form, the 2D projective transformation,

$$\forall \boldsymbol{x}_i \leftrightarrow \boldsymbol{x}_i^{'} \in \mathcal{S}, \quad c_i \boldsymbol{x}_i^{'} = \boldsymbol{H}_{3\times 3}\boldsymbol{x}_i$$

*Since all the keypoints are normal points (instead of infinity points), we can suppose* $\boldsymbol{x}_i$ *and* $\boldsymbol{x}_i^{'}$ *are of the form "normalized homogeneous coordinate"*

$$c_i \begin{bmatrix} x_i^{'} \\ y_i^{'} \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

# Homography estimation—Modeled as a homogeneous linear LS problem

Our task: solving ***H*** from

$$c_i\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, i = 1, 2, ..., p \qquad (1)$$

$$\begin{cases} h_{11}x_i + h_{12}y_i + h_{13} = c_i x_i' \\ h_{21}x_i + h_{22}y_i + h_{23} = c_i y_i' \\ h_{31}x_i + h_{32}y_i + h_{33} = c_i \end{cases} \Rightarrow \begin{cases} \dfrac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} = x_i' \\ \dfrac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} = y_i' \end{cases} \Rightarrow \begin{pmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' & -y_i' \end{pmatrix}\begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = \mathbf{0}$$

If $p$=4, we have,

$$\boldsymbol{A}_{8\times 9}\boldsymbol{h}_{9\times 1} = \mathbf{0} \qquad (2)$$

Normally, rank($\boldsymbol{A}$)=8; thus (2) has 1 linear independent solution vector in its solution space

**4 correspondence pairs can uniquely determine a 2D projective transformation (Homography)**

# Homography estimation—Modeled as a homogeneous linear LS problem

Our task: solving $\boldsymbol{H}$ from

$$c_i \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, i = 1, 2, ..., p \qquad (1)$$

$$\begin{cases} h_{11}x_i + h_{12}y_i + h_{13} = c_i x_i' \\ h_{21}x_i + h_{22}y_i + h_{23} = c_i y_i' \\ h_{31}x_i + h_{32}y_i + h_{33} = c_i \end{cases} \Rightarrow \begin{cases} \dfrac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} = x_i' \\ \dfrac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} = y_i' \end{cases} \Rightarrow \begin{pmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' & -x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' & -y_i' \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = \mathbf{0}$$

If $p$=4, we have,

$$\boldsymbol{A}_{8\times 9}\boldsymbol{h}_{9\times 1} = \mathbf{0} \qquad (2)$$

Normally, rank($\boldsymbol{A}$)=8; thus (2) has 1 linear independent solution vector in its solution space

**However, in most cases $p$>4** $\Rightarrow$

# Homography estimation—Modeled as a homogeneous linear LS problem

$$\boldsymbol{A}_{2p\times 9}\boldsymbol{h}_{9\times 1} = \boldsymbol{0},\ p > 4 \tag{3}$$

Normally, rank($A$)=9; thus (3) **only has a zero solution**, which is of no use for us!

- Notice two facts
  - What we want is a non-zero solution $\boldsymbol{h}^*$ that can **roughly** satisfy Eq. (3)
  - $\boldsymbol{h}^*$ actually represents a projective transformation; and $\forall k \neq 0, k\boldsymbol{h}^*$ represents the same projective transformation with $\boldsymbol{h}^*$. Without loss of generality, we can require that $\left\|\boldsymbol{h}^*\right\|_2^2 = 1$

Thus, our problem can be formulated as,

$$\boldsymbol{h}^* = \arg\min_{\boldsymbol{h}} \left\|\boldsymbol{A}\boldsymbol{h}\right\|_2^2, \text{subject to} \left\|\boldsymbol{h}\right\|_2^2 = 1, \boldsymbol{A} \in \mathbb{R}^{2p\times 9}, \boldsymbol{h} \in \mathbb{R}^{9\times 1}, rank(\boldsymbol{A}) = 9$$

A more general form (**the homogeneous linear least-squares problem**),

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \left\|\boldsymbol{A}\boldsymbol{x}\right\|_2^2, \text{subject to} \left\|\boldsymbol{x}\right\|_2^2 = 1, \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}, rank(\boldsymbol{A}) = n \tag{4}$$

➡ Let's solve it!

# Homography estimation—Modeled as a homogeneous linear LS problem

Construct the Lagrange function,

$$L(\boldsymbol{x},\lambda)=\|\boldsymbol{A}\boldsymbol{x}\|_2^2+\lambda\left(1-\|\boldsymbol{x}\|_2^2\right)$$

Solving the stationary points of $L(\boldsymbol{x},\lambda)$. $(\boldsymbol{x}_0,\lambda_0)$ is a stationary of $L(\boldsymbol{x},\lambda)$, if only if,

$$\begin{cases}\dfrac{\partial L}{\partial \boldsymbol{x}}\Big|_{\boldsymbol{x}=\boldsymbol{x}_0,\lambda=\lambda_0}=\mathbf{0}\\ \dfrac{\partial L}{\partial \lambda}\Big|_{\boldsymbol{x}=\boldsymbol{x}_0,\lambda=\lambda_0}=0\end{cases}\Rightarrow\begin{cases}\dfrac{\partial\left(\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}+\lambda\left(1-\boldsymbol{x}^T\boldsymbol{x}\right)\right)}{\partial \boldsymbol{x}}\Big|_{\boldsymbol{x}=\boldsymbol{x}_0,\lambda=\lambda_0}=\mathbf{0}\\ \dfrac{\partial\left(\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}+\lambda\left(1-\boldsymbol{x}^T\boldsymbol{x}\right)\right)}{\partial \lambda}\Big|_{\boldsymbol{x}=\boldsymbol{x}_0,\lambda=\lambda_0}=0\end{cases}\Rightarrow\begin{cases}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}_0=\lambda_0\boldsymbol{x}_0\\ \boldsymbol{x}_0^T\boldsymbol{x}_0=1\end{cases}$$

Note: the stationary point of $L(\boldsymbol{x},\lambda)$ is not unique

Suppose that $(\boldsymbol{x}_i,\lambda_i)$ is a stationary point of $L(\boldsymbol{x},\lambda)$, then $\boldsymbol{x}_i$ is a possible extremum point of $f(\boldsymbol{x})=\|\boldsymbol{A}\boldsymbol{x}\|_2^2$ under the equality constraint and we have,

$$f(\boldsymbol{x}_i)=\|\boldsymbol{A}\boldsymbol{x}_i\|_2^2=\boldsymbol{x}_i^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}_i=\boldsymbol{x}_i^T\lambda_i\boldsymbol{x}_i=\lambda_i$$

Thus, the minimum of $f(\boldsymbol{x})$ (under the equality constraint) is $\min\{\lambda_i\}_{i=1}^n$, that is, the least eigen-value of $\boldsymbol{A}^T\boldsymbol{A}$. And the minimizer of $f(\boldsymbol{x})$ is the unit eigen-vector of $\boldsymbol{A}^T\boldsymbol{A}$ associated with its least eigen value

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Problem formulation

We know that a $3\times3$ matrix $\boldsymbol{H}$, representing a valid 2D projective transform, has 8 DoF. Thus, when optimizing a $9d$ vector representing a projective transform, we need to constrain its DoF to 8. When the problem is modeled as a homogeneous linear LS problem, such a goal is achieved by imposing a constraint $\|\boldsymbol{h}\|_2^2=1$

In practical applications, another method can be used to constrain the DoF of $\boldsymbol{H}$ to 8, that is by fixing an element of $\boldsymbol{H}$ to a non-zero constant (such an element in the ground-truth $\boldsymbol{H}$ should be non-zero). Usually, we can fix $h_{33}=1$

$\Downarrow$

Our task: solving $\boldsymbol{H}$ from

$$c_i\begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & \boxed{h_{33}} \end{bmatrix}\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, i=1,2,\ldots,p \tag{1}$$

fixed to $1$

$\Downarrow$

# Problem formulation

$$c_i \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, i = 1, 2, ..., p \tag{5}$$

$$\begin{cases} h_{11}x_i + h_{12}y_i + h_{13} = c_i x_i' \\ h_{21}x_i + h_{22}y_i + h_{23} = c_i y_i' \\ h_{31}x_i + h_{32}y_i + 1 = c_i \end{cases} \Rightarrow \begin{cases} \dfrac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + 1} = x_i' \\ \dfrac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + 1} = y_i' \end{cases} \Rightarrow \begin{pmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_i x_i' & -y_i x_i' \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i y_i' & -y_i y_i' \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} x_i' \\ y_i' \end{pmatrix}$$

If $p$=4, we have,

$$\boldsymbol{A}_{8\times 8}\boldsymbol{h}_{8\times 1} = \boldsymbol{b}_{8\times 1} \tag{6}$$

Normally, rank($\boldsymbol{A}$)=rank([$\boldsymbol{A}$; $\boldsymbol{b}$])=8; thus (6) has 1 unique solution

4 correspondence pairs can uniquely determine a 2D projective transformation (Homography)

# Problem formulation

$$c_i\begin{bmatrix}x_i'\\y_i'\\1\end{bmatrix}=\begin{bmatrix}h_{11}&h_{12}&h_{13}\\h_{21}&h_{22}&h_{23}\\h_{31}&h_{32}&1\end{bmatrix}\begin{bmatrix}x_i\\y_i\\1\end{bmatrix}, i=1,2,...,p \qquad (5)$$

$$\begin{cases}h_{11}x_i+h_{12}y_i+h_{13}=c_ix_i'\\h_{21}x_i+h_{22}y_i+h_{23}=c_iy_i'\\h_{31}x_i+h_{32}y_i+1=c_i\end{cases} \Rightarrow \begin{cases}\dfrac{h_{11}x_i+h_{12}y_i+h_{13}}{h_{31}x_i+h_{32}y_i+1}=x_i'\\\dfrac{h_{21}x_i+h_{22}y_i+h_{23}}{h_{31}x_i+h_{32}y_i+1}=y_i'\end{cases} \Rightarrow \begin{pmatrix}x_i&y_i&1&0&0&0&-x_ix_i'&-y_ix_i'\\0&0&0&x_i&y_i&1&-x_iy_i'&-y_iy_i'\end{pmatrix}\begin{pmatrix}h_{11}\\h_{12}\\h_{13}\\h_{21}\\h_{22}\\h_{23}\\h_{31}\\h_{32}\end{pmatrix}=\begin{pmatrix}x_i'\\y_i'\end{pmatrix}$$

If $p$=4, we have,

$$\boldsymbol{A}_{8\times8}\boldsymbol{h}_{8\times1}=\boldsymbol{b}_{8\times1} \qquad (6)$$

Normally, rank($\boldsymbol{A}$)=rank([$\boldsymbol{A}$; $\boldsymbol{b}$])=8; thus (6) has 1 unique solution

**However, in most cases $p$>4** $\Rightarrow$

## Problem formulation

$$\boldsymbol{A}_{2p\times 8}\boldsymbol{h}_{8\times 1}=\boldsymbol{b}_{2p\times 1},\ p>4 \qquad (7)$$

Normally, $rank(\boldsymbol{A})=8, rank([\boldsymbol{A};\boldsymbol{b}])=9, rank(\boldsymbol{A})\neq rank([\boldsymbol{A};\boldsymbol{b}])$, and thus Eq. 7 does not have a solution!

⬇

We want a solution that can *roughly* satisfy Eq. 7

⬇

Thus, our problem can be formulated as,

$$\boldsymbol{h}^*=\arg\min_{\boldsymbol{h}}\|\boldsymbol{A}\boldsymbol{h}-\boldsymbol{b}\|_2^2,\ \boldsymbol{A}\in\mathbb{R}^{2p\times 8}, \boldsymbol{h}\in\mathbb{R}^{8\times 1}, \boldsymbol{b}\neq\boldsymbol{0}\in\mathbb{R}^{2p\times 1}$$

where $rank(\boldsymbol{A})=8$

A more general form (**the inhomogeneous linear least-squares problem**),

$$\boldsymbol{x}^*=\arg\min_{\boldsymbol{x}}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2,\ \boldsymbol{A}\in\mathbb{R}^{m\times n}, \boldsymbol{x}\in\mathbb{R}^{n\times 1}, \boldsymbol{b}\neq\boldsymbol{0}\in\mathbb{R}^{m\times 1}, rank(\boldsymbol{A})=n \qquad (8)$$

➡ Let's solve it!

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

## Method 1 (requires that $\boldsymbol{A}$ has a full column rank)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2, \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}, \boldsymbol{b} \neq \boldsymbol{0} \in \mathbb{R}^{m\times 1}, rank(\boldsymbol{A}) = n \qquad (8)$$

| Exercise | Please prove that the objective function in (8), $f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2, \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}, \boldsymbol{b} \neq \boldsymbol{0} \in \mathbb{R}^{m\times 1}$ is a convex function |
|---|---|

The stationary point of the objective function $f(\boldsymbol{x})$ is $\boldsymbol{x}_s = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b}$

Since $f(\boldsymbol{x})$ is convex

$$\boldsymbol{x}^* = \boldsymbol{x}_s = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b} \qquad (9)$$

**(For a convex function, its stationary point is the global minimizer[1]. We will discuss more about the convex optimization in Chapter 13)**

[1] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004, pp. 69

# Method 1 (requires that $\boldsymbol{A}$ has a full column rank)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2, \boldsymbol{A}\in\mathbb{R}^{m\times n}, \boldsymbol{x}\in\mathbb{R}^{n\times 1}, \boldsymbol{b}\neq\boldsymbol{0}\in\mathbb{R}^{m\times 1}, rank(\boldsymbol{A})=n \quad (8)$$

**Exercise**

Please prove that the objective function in (8),

$$f(\boldsymbol{x}) = \|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2, \boldsymbol{A}\in\mathbb{R}^{m\times n}, \boldsymbol{x}\in\mathbb{R}^{n\times 1}, \boldsymbol{b}\neq\boldsymbol{0}\in\mathbb{R}^{m\times 1}$$

is a convex function

The stationary point of the objective function $f(\boldsymbol{x})$ is $\boldsymbol{x}_s = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b}$

Is it really invertible?

Since $f(\boldsymbol{x})$ is convex

$$\boldsymbol{x}^* = \boldsymbol{x}_s = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b} \quad (9)$$

**Assignment**

For the matrix $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, if $rank(\boldsymbol{A})=n$, $\boldsymbol{A}^T\boldsymbol{A}$ should be invertible

The "method 1" requires that $\boldsymbol{A}$ has a full column rank. Next, we will learn a more general and accurate method to solve the inhomogeneous linear LS problem

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

## Method 2—A more general and accurate approach

- For solving the linear least squares numerically with a computer, usually we do not use the form of Eq. (9) (though it is elegant) for two reasons
  - When rank($\boldsymbol{A}$)<$n$, $\boldsymbol{x}^*$ can not be determined
  - Even though $\boldsymbol{A}^T\boldsymbol{A}$ is invertible, the formation of $\boldsymbol{A}^T\boldsymbol{A}$ can dramatically degrade the accuracy of the computation
- Instead, we can use the technique of SVD

Our problem,

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \left\| \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \right\|_2^2, \boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{x} \in \mathbb{R}^{n\times 1}, \boldsymbol{b} \neq \boldsymbol{0} \in \mathbb{R}^{m\times 1} \quad (10)$$

**Note that this problem is more general than Eq. 8 since its does not have additional requirements on the rank of *A***

## Method 2—A more general and accurate approach

Suppose the SVD form of $\boldsymbol{A}$ is,

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

$$\Downarrow$$

$$\boldsymbol{Ax} - \boldsymbol{b} = \boldsymbol{U}\Sigma\boldsymbol{V}^{T}\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{U}\left(\Sigma\boldsymbol{V}^{T}\boldsymbol{x}\right) - \boldsymbol{U}\left(\boldsymbol{U}^{T}\boldsymbol{b}\right) \triangleq \boldsymbol{U}\left(\Sigma\boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1}\right)$$

where $\boldsymbol{y}_{n\times 1} = \boldsymbol{V}^{T}\boldsymbol{x}, \boldsymbol{c}_{m\times 1} = \boldsymbol{U}^{T}\boldsymbol{b}$

Since $\boldsymbol{U}$ is an orthogonal matrix,

$$\|\boldsymbol{Ax} - \boldsymbol{b}\| = \left\|\boldsymbol{U}\left(\Sigma\boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1}\right)\right\| = \left\|\Sigma\boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1}\right\|$$

Then, our objective is to identify $\boldsymbol{y}$ that can make $\left\|\Sigma\boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1}\right\|$ have the minimum length

# Method 2—A more general and accurate approach

$$\Sigma \boldsymbol{y}_{n\times 1} = \begin{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix} & \boldsymbol{O}_{r\times(n-r)} \\ \boldsymbol{O}_{(m-r)\times r} & \boldsymbol{O}_{(m-r)\times(n-r)} \end{bmatrix}_{m\times n} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sigma_1 y_1 \\ \sigma_2 y_2 \\ \vdots \\ \sigma_r y_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{m\times 1} \Rightarrow \Sigma \boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1} = \begin{bmatrix} \sigma_1 y_1 - c_1 \\ \sigma_2 y_2 - c_2 \\ \vdots \\ \sigma_r y_r - c_r \\ -c_{r+1} \\ \vdots \\ -c_m \end{bmatrix}_{m\times 1}$$

Then, we simply let $y_i = \frac{c_i}{\sigma_i}, 1 \le i \le r$ ; then, $\left\| \Sigma \boldsymbol{y}_{n\times 1} - \boldsymbol{c}_{m\times 1} \right\|$ can get the minimum length $\sqrt{\sum_{i=r+1}^{m} c_i^2}$

Note that $y_{r+1} \sim y_n$ can be arbitrary

# Method 2—A more general and accurate approach

The operation $y_i = \dfrac{c_i}{\sigma_i}, 1 \le i \le r$ can be simply completed by a matrix multiplication,

$$\boldsymbol{y} = \begin{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1} & & & \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_r} \end{bmatrix} & \boldsymbol{O}_{r\times(m-r)} \\ \boldsymbol{O}_{(n-r)\times r} & \boldsymbol{O}_{(n-r)\times(m-r)} \end{bmatrix}_{n\times m} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}_{m\times 1} = \begin{bmatrix} c_1/\sigma_1 \\ c_2/\sigma_2 \\ \vdots \\ c_r/\sigma_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1} \triangleq \boldsymbol{\Sigma}^+ \boldsymbol{c}_{m\times 1}$$

where $\Sigma^+$ means transposing $\Sigma$ and inverting all non-zero diagonal entries

Finally,

$$\boldsymbol{x} = \boldsymbol{V}\boldsymbol{y}_{n\times 1} = \boldsymbol{V}\boldsymbol{\Sigma}^+\boldsymbol{c}_{m\times 1} = \boldsymbol{V}\boldsymbol{\Sigma}^+\boldsymbol{U}^T\boldsymbol{b} \qquad (11)$$

Moore-Penrose inverse

# Method 2—A more general and accurate approach

**Roger Penrose** (born 8 August 1931) is an English mathematician, mathematical physicist, philosopher of science and Nobel Laureate in Physics. He is Emeritus Rouse Ball Professor of Mathematics in the University of Oxford.

Penrose has contributed to the mathematical physics of general relativity and cosmology. He has received several prizes and awards, including the 1988 Wolf Prize in Physics, which he shared with Stephen Hawking for the **Penrose–Hawking singularity theorems**, and the 2020 Nobel Prize in Physics for the discovery that black hole formation is a robust prediction of the general theory of relativity.

# Method 2—A more general and accurate approach

- Some notes about the Moore-Penrose inverse used in linear least squares
  - It does not have requirements for the rank of $\boldsymbol{A}$
  - It can guarantee that the obtained solution can make $\|\boldsymbol{Ax}-\boldsymbol{b}\|$ having the minimum length; but **the solution may be not unique**

# Outline

- Prerequisites
  - Matrix Differentiation
  - Lagrange Multiplier
  - Singular Value Decomposition
- Homography estimation—Modeled as a homogeneous linear LS problem
- Homography estimation—Modeled as an inhomogeneous linear LS problem
  - Problem Formulation
  - Method 1 (requires that $A$ has a full column rank)
  - Method 2—A more general and accurate approach
- Summary of solvers for the linear LS problem

# Summary of solvers for the linear LS problem

To solve the equation,

$$\boldsymbol{Ax}=\boldsymbol{b}, \boldsymbol{A}\in\mathbb{R}^{m\times n}, \boldsymbol{x}\in\mathbb{R}^{n\times 1}, \boldsymbol{b}\in\mathbb{R}^{m\times 1}$$

if $\boldsymbol{b}=\boldsymbol{0}$

$$\boldsymbol{x}^*=\arg\min_{\boldsymbol{x}}\|\boldsymbol{Ax}\|_2^2, \text{subject to } \|\boldsymbol{x}\|_2^2=1$$

$\boldsymbol{x}^*$ is the unit eigen-vector of $\boldsymbol{A}^T\boldsymbol{A}$ associated with its least eigen value

if $\boldsymbol{b}!=\boldsymbol{0}$

$$\boldsymbol{x}^*=\arg\min_{\boldsymbol{x}}\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2, \boldsymbol{b}\neq\boldsymbol{0}\in\mathbb{R}^{m\times 1}$$

if rank($\boldsymbol{A}$)=$n$

$$\boldsymbol{x}^*=\boldsymbol{x}_s=\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

If we do not know the rank of $\boldsymbol{A}$

$$\boldsymbol{x}^*=\boldsymbol{V}\boldsymbol{\Sigma}^+\boldsymbol{U}^T\boldsymbol{b}$$


Q+A