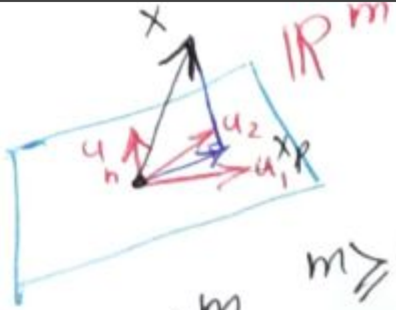
Mathematics for AI

Lecture 8
Orthonormal Basis, QR decomposition, least squares solution

Remember: Orthogonal Projection



$m \geq n$

$x \in \mathbb{R}^m$

$a \in \mathbb{R}^n$

$$\underset{m\times 1}{x_p} = \underset{m\times n}{\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}} \underset{n\times 1}{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}}$$

$$x_p = U\vec{a}$$

$$\vec{a} = (U^TU)^{-1}U^T x$$

$$\underset{m\times 1}{x_p} = \underbrace{\underset{m\times n}{U}\,\underset{n\times n}{(U^TU)^{-1}}\,\underset{n\times m}{U^T}}_{m\times m}\,\underset{m\times 1}{x}$$

Orthonormal Basis

The basis vectors $u_1, u_2 \cdots u_n$ are unit vectors &

~~orth~~ perpendicular to each other

$$\|u_i\| = 1 \Longrightarrow u_i^T u_j = \langle u_i, u_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$\langle u_i, u_j \rangle = 0 \quad i \neq j$$

$\Longrightarrow$ $u_1, u_2, \ldots, u_n$ are orthonormal vectors.

$$\begin{array}{cc} m & \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \\ & n \end{array} = U$$

$$U^T U = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_n^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$$

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

$$(U^T U)_{ij} = u_i^T u_j$$

$$U^T U = I$$

Orthonormal Basis

MA8

$$U = \begin{array}{c c} & n \\ m & \begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ & & & \end{bmatrix} \end{array}$$

$n \le m$

$\mathrm{rank}(U) = n$

$u_1, u_2, \ldots, u_n$ are linearly independent

$$\mathrm{rank}(AB) \le \min(\mathrm{rank}(A), \mathrm{rank}(B))$$

$$U^T U = \underbrace{\begin{bmatrix} U^T \end{bmatrix}}_{n \times m} \underbrace{\begin{bmatrix} U \end{bmatrix}}_{m \times n} = \underbrace{\begin{bmatrix} U^T U \end{bmatrix}}_{n \times n}$$

$$U U^T = \underbrace{\begin{bmatrix} U \end{bmatrix}}_{m \times n} \underbrace{\begin{bmatrix} U^T \end{bmatrix}}_{n \times m} = \underbrace{\begin{bmatrix} U U^T \end{bmatrix}}_{m \times m}$$

$$= \underbrace{u_1 u_1^T}_{m \times m} + \underbrace{u_2 u_2^T}_{m \times m} + \cdots + \underbrace{u_n u_n^T}_{m \times m}$$

Matrix with orthonormal columns

if $U \in \mathbb{R}^{m \times n}$ has orthonormal columns $U^T U = I$

rank(U)

I $n < m$ $\quad UU^T \neq I \quad [U]$

Orthogonal Matrix

II $m = n$ U square $(n \times n)$ $U^T U = I$

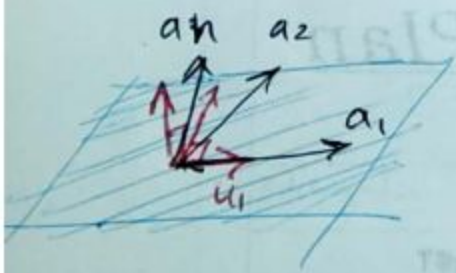
$\text{rank}(U) = n$ $U^T = U^{-1}$

$U U^T = I$

If ~~A~~ U is square and its columns ~~a~~ are orthonormal then its rows are also orthonormal.

$$\left.\begin{array}{l} U^T U = I = U U^T \\ U \in \mathbb{R}^{n \times n} \text{ (sqare)} \end{array}\right\} \text{orthogonal matrix}$$

Gram-Scmit Orthogonalization



$a_1, a_2 \ldots a_n \in \mathbb{R}^n$ for ~~an orthonormal~~

a basis for $\mathbb{R}^n$

$$u_1 = \frac{a_1}{\|a_1\|}$$

$$\tilde{u}_2 = a_2 - u_1 u_1^T a_2 = (I - u_1 u_1^T) a_2$$

Gram-Scmit orthogonalization

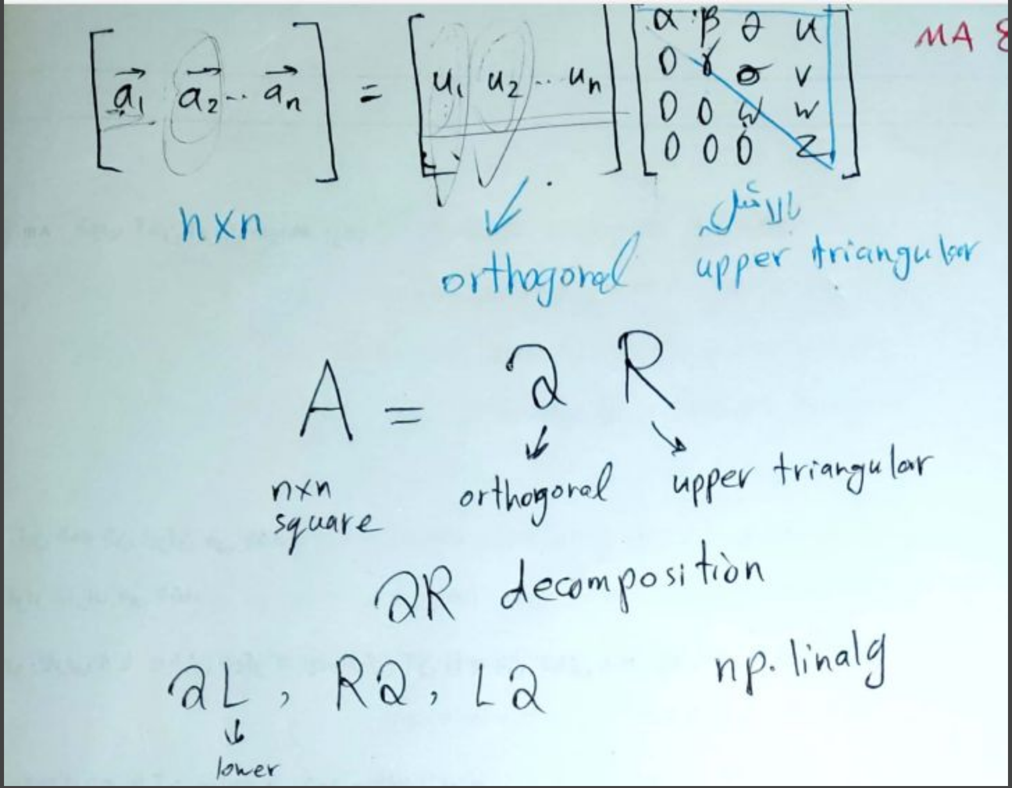
$$u_2 = \frac{u_2}{\|u_2\|}$$

$$\tilde{u}_3 = a_3 - \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} a_3$$

$$= a_3 - u_1 u_1^T a_3 - u_2 u_2^T a_3$$

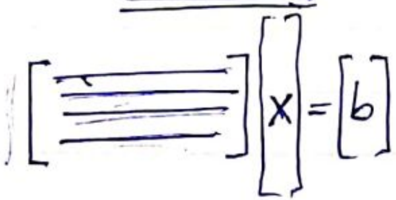
QR decomposition

K. N. Toosi
University of Technology



Underdetermined system

$$Ax=b$$



$$\begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

$A \in \mathbb{R}^{m \times n}$

$r = \text{Rank}(A) < n$

$b \in C(A)$ $\Rightarrow$

$r < n$ independen equations

n unknown

Underdetermined Equations

Overdetermined system

K. N. Toosi
University of Technology

$$Ax = b$$

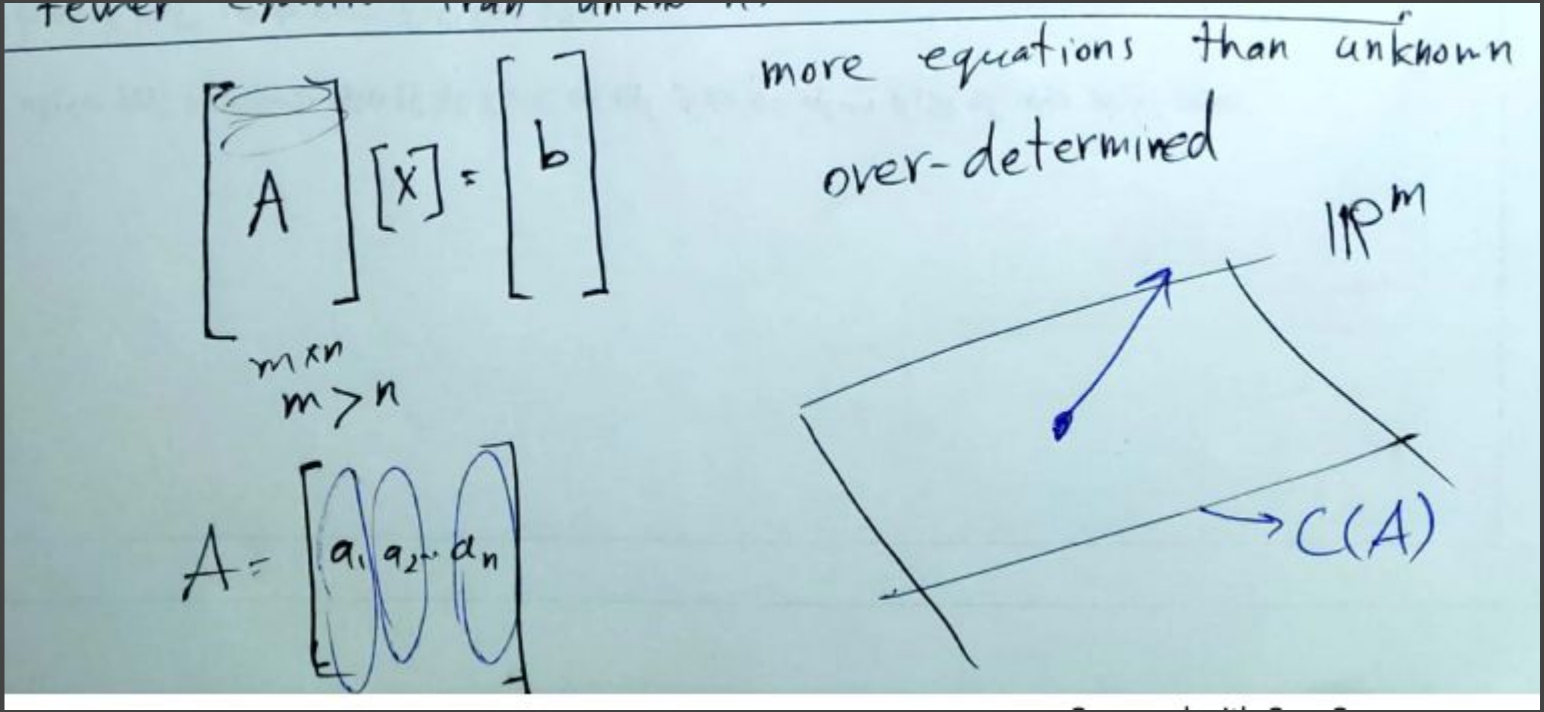
$$\begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \\ \\ \\ \end{bmatrix}$$

$$A \in \mathbb{R}^{m \times n}$$

$$b \notin C(A)$$

<u>Overdetermined</u>

Overdetermined system



Linear measurements

$$y_1 = a_{11} x_1 + a_{12} x_2 + a_{13} x_3$$

$$y_2 = a_{21} x_1 + a_{22} x_2 + a_{23} x_3$$

$$y_3 = a_{31} x_1 + a_{32} x_2 + a_{33} x_3$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

measurement

Noisy linear measurements

K. N. Toosi
University of Technology

$y_1 = a_{11} x_1 + a_{12} x_2 + a_{13} x_3 + n_1 \rightarrow$ (measument process) noise

$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$y_2 = a_{21} x_1 + a_{22} x_2 + a_{23} x_3 + n_2$

$y_3 = a_{31} x_1 + a_{32} x_2 + a_{33} x_3 + n_3$

$y = Ax + n$ noise

$y_4 = a_{41} x_1 + a_{42} x_2 + a_{43} x_3 + n_4$

$n \in \mathbb{R}^m$

$y_5 =$

$y_6 =$

II

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} + \begin{bmatrix} n \end{bmatrix}$$

noise is unknown

noise is (relatively) small

$y \notin C(A)$

$y - n \in C(A)$

$\boxed{Ax = y} \rightarrow$ overdetermind

$A \in \mathbb{R}^{m \times n}$

What is the best solution?

K. N. Toosi
University of Technology

$y \notin C(A)$

$y - n \in C(A)$

$\boxed{Ax = y} \rightarrow$ overdetermind

$A \in \mathbb{R}^{m \times n}$

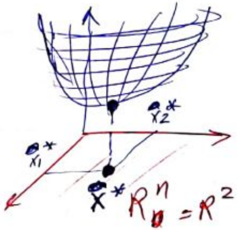
there is no x such that $A\underline{x} = y$

find an x such that $\|Ax - y\|$ is small

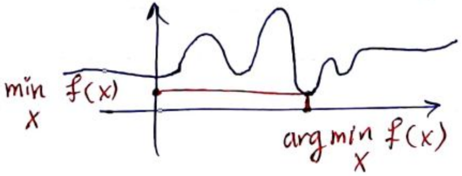
$\|u\|$ = lenght of u.

$Ax - y \in \mathbb{R}^m$

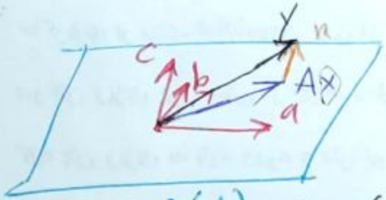
$f(x) = \|Ax - y\|$ $\quad f: \mathbb{R}^n \rightarrow \mathbb{R}$



$x^* = \underset{x}{\operatorname{argmin}}\ f(x)$



What is the best solution?



$C(A) = \text{span}(\vec{a}, \vec{b}, \vec{c})$

$Ax = y - n$ $\quad n = y - Ax$ مجهول در کوچک

$y \notin A$ $\quad Ax = y$ has no solution

find x such ~~to~~ that $\|Ax - y\|$ is small.

$$x = \arg\min_{x} \|Ax - y\|$$

$2 = \min_{x} f(x)$ $\quad 6 = \arg\min_{x} f(x)$

What is the best solution?

K. N. Toosi
University of Technology

$$\begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} + \begin{bmatrix} n \end{bmatrix}$$

$Ax = y$ overdetermined

III

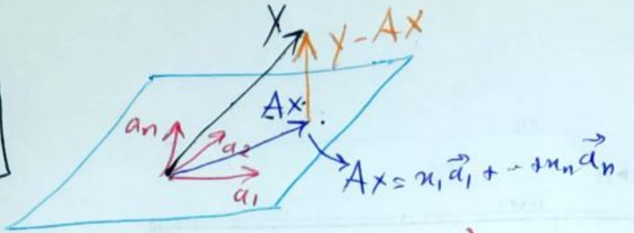
$$x^* = \underset{x}{\operatorname{argmin}} \; \| Ax - y \|$$

Least Squares

K. N. Toosi
University of Technology

$$x = \arg\min_{x} \|Ax - y\|$$

$$Ax = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$



$$C(A) = \mathrm{span}(a_1, a_2, \dots, a_n)$$

best x is when Ax becoms the projection of y into C(A). $\Rightarrow \begin{cases} y - Ax \perp a_1 \\ y - Ax \perp a_2 \\ \vdots \\ y - Ax \perp a_n \end{cases}$

$$y_p = Ax = A(A^TA)^{-1}A^Ty$$

$$x = (A^TA)^{-1}A^Ty = \arg\min_{x} \|Ax - y\|$$

Least Squares

$y = Ax \perp a_n$

$y_p = Ax = A\underline{(A^TA)^{-1}A^Ty}$ مهم است

$x = (A^TA)^{-1}A^Ty = \underset{x}{\operatorname{argmin}} \|Ax - y\|$

$Ax \simeq y \qquad x = (A^TA)^{-1}A^Ty$

↳ least squares solution

جواب کمینه مربعات

$\underset{x}{\operatorname{argmin}} \|Ax - y\| = \underset{x}{\operatorname{argmin}} \|Ax - y\|^2$

$$\rho = Ax - y = \begin{bmatrix} r_1^T \\ r_2^T \\ \vdots \\ r_m^T \end{bmatrix} x - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} r_1^T x \\ r_2^T x \\ \vdots \\ r_m^T x \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} r_1^T x - y_1 \\ r_2^T x - y_2 \\ \vdots \\ r_m^T x - y_m \end{bmatrix}$$

↓ residual vector

$$\|\rho\|^2 = \left\| \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_m \end{bmatrix} \right\|^2 = \sum_{i=1}^{m} \rho_i^2 = \sum_{i=1}^{m} \left(r_i^T x - y_i\right)^2$$

x = جواب کمینه مربعات . مجموع مربعات را کمینه می کند.

Least Squares Solution

K. N. Toosi
University of Technology

$Ax = y$ A has full column rank

IV
16

$$\begin{bmatrix} A \end{bmatrix} x = \begin{bmatrix} y \end{bmatrix} = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

$$x^* = \arg\min \|Ax - y\|^2 = \arg\min_x \left\| \begin{bmatrix} a_1^T x - y_1 \\ a_2^T x - y_2 \\ \vdots \\ a_m^T x - y_m \end{bmatrix} \right\|^2$$

$$= \underbrace{\arg\min_x \underbrace{(a_1^T x - y_1)^2 + (a_2^T x - y_2)^2 + \cdots + (a_m^T x - y_m)^2}_{\text{sum of squares}\quad مجموع\ مربعات}}_{\text{Least Squares}}$$

Least Squares مسأله کمینه (مجموع) مربعات

x^*: Least Squares Solution

b
$e = b - Ax^*$

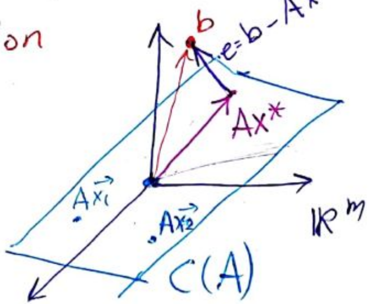
Geometric Interpretation

Least Squares مسأله کمینه (مجموع) مربعات

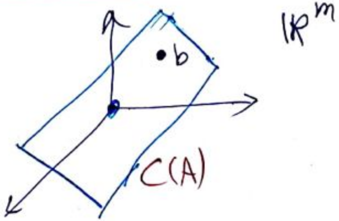
x^*: Least Squares Solution

$Ax = b \qquad b \notin C(A)$

$\min_x \|Ax - b\| = \mathrm{dist}(b, C(A))$



$\boxed{b \in C(A) \Rightarrow \min_x \|Ax - b\| = 0}$

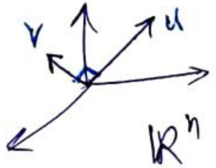


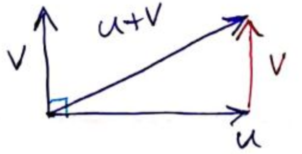
$\Rightarrow (b - Ax^*) \perp Ax^*$

$(b - Ax^*) \perp Ax$ for all x

Orthogonal vectors

K. N. Toosi
University of Technology





قضیه فیثاغورث :

$$u \perp v \Rightarrow \|u\|^2 + \|v\|^2 = \|u+v\|^2$$

$$u^T u + v^T v = (u+v)^T (u+v)$$

$$= (u^T + v^T)(u+v)$$

$$= u^T u + \underbrace{u^T v + v^T u}_{2u^Tv} + v^T v \in \mathbb{R}$$

$$\Rightarrow \quad u \perp v \Leftrightarrow \cancel{u^T u} + \cancel{v^T v} = \cancel{u^T u} + 2u^T v + \cancel{v^T v}$$

$$\Rightarrow 2u^T v = 0 \Rightarrow u^T v = 0 = \langle u, v \rangle$$

$$\boxed{u \perp v \Leftrightarrow \langle u, v \rangle = 0}$$