# Numerical Problems in Solving the Normal Equation

Note Title

In general, it is not a good idea to solve the normal eqn:

$$A^T A \mathbb{x} = A^T \mathbb{b}$$

by explicitly forming $A^T A$, and then compute $(A^T A)^{-1}$!

<u>why?</u>

1) Forming $A^T A \to$ loss of info.
2) $\kappa(A^T A) = \kappa(A)^2$, i.e.,

the cond. number of $A^T A$ is much worse than that of $A$ in general.

This example is a bit extreme ... Show previous MATLAB example

<u>Ex.</u> <u>Forming $A^T A$ is bad.</u>

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \quad \text{say } \varepsilon = 10^{-8}$$

in double precision floating point sys.

Then $$A^T A = \begin{bmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{bmatrix}$$

$$\approx \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$ because $\varepsilon^2 = 10^{-16}$

<u>How about the condition numbers?</u>

$\kappa(A) \approx 1.4142 \times 10^8$ already bad.

$\kappa(A^T A) \approx +\infty$ in double precision.

If we set $\varepsilon = 10^{-7}$ instead of $10^{-8}$,
then $\kappa(A) \approx 1.4142 \times 10^{7}$
$\kappa(A^T A) \approx 1.9903 \times 10^{14}$
This is still too bad to get any reliable LS solution for such $A$.

Often such situations occur when some of the column vectors of $A$ are "close to parallel", i.e., they become almost linearly dependent.

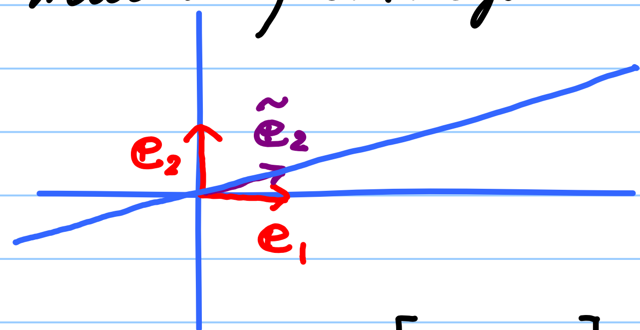
Def. Let $A \in \mathbb{R}^{m \times n}$. Then
$A$ is called rank deficient if
$\mathrm{rank}(A) < \min(m, n)$.

i.e., if $A$ is not of full rank.

In general, we should avoid computing a solution for a given LS problem by forming $A^T A$ explicitly and computing $(A^T A)^{-1} A^T \mathbf{b}$.
$\Rightarrow$ Better to use the methods based on QR decomposition or SVD (We'll discuss these later in this course.)

# Orthogonality

The above discussion should convince you that $A$ is quite "good" if its column vectors are mutually orthogonal.




Suppose $A = [\mathbf{e}_1 \ \mathbf{e}_2]$, $\tilde{A} = [\mathbf{e}_1 \ \tilde{\mathbf{e}}_2]$ in $\mathbb{R}^2$. You can see that $A$ is much more "well-balanced" and convenient than $\tilde{A}$. For example, suppose we want to represent $\mathbf{x} = [1, 1]^T$ in the basis of $\{\mathbf{e}_1, \mathbf{e}_2\}$ and that of $\{\mathbf{e}_1, \tilde{\mathbf{e}}_2\}$. Then the coefficient of $\mathbf{x}$ w.r.t. $\{\mathbf{e}_1, \mathbf{e}_2\}$ is the same as $\mathbf{x}$ itself since $A^{-1}\mathbf{x} = A\mathbf{x} = \mathbf{x}$

$A = I$ in $\mathbb{R}^2$

But $\tilde{A}^{-1}\mathbf{x}$ behaves badly.

Why? Say $\mathbf{c} = \tilde{A}^{-1}\mathbf{x}$, $\mathbf{c} = [c_1, c_2]^T$

Then $\mathbf{x} = \tilde{A}\mathbf{c} = [\mathbf{e}_1 \ \tilde{\mathbf{e}}_2]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$$= c_1\mathbf{e}_1 + c_2\tilde{\mathbf{e}}_2$$

But $\mathbf{x} = \mathbf{e}_1 + \mathbf{e}_2$, i.e.,

$$\mathbf{e}_1 + \mathbf{e}_2 = c_1\mathbf{e}_1 + c_2\tilde{\mathbf{e}}_2$$

Taking an inner product with $\mathbf{e}_2$ on both sides yields

$$\underbrace{\mathbf{e}_2^T(\mathbf{e}_1+\mathbf{e}_2)}_{\substack{\| \\ \mathbf{e}_2^T\mathbf{e}_2 \\ \| \\ \|\mathbf{e}_2\|_2^2 = 1.}} = \underbrace{\mathbf{e}_2^T(c_1\mathbf{e}_1 + c_2\tilde{\mathbf{e}}_2)}_{\substack{\| \\ c_1 \underbrace{\mathbf{e}_2^T\mathbf{e}_1}_{=0} + c_2\mathbf{e}_2^T\tilde{\mathbf{e}}_2 \\ \| \\ c_2\mathbf{e}_2^T\tilde{\mathbf{e}}_2}}$$

$$\Rightarrow \quad 1 = c_2\mathbf{e}_2^T\tilde{\mathbf{e}}_2$$

$$\Rightarrow \quad c_2 = \frac{1}{\mathbf{e}_2^T\tilde{\mathbf{e}}_2}$$

could be huge if $\tilde{\mathbf{e}}_2$ is close to perpendicular to $\mathbf{e}_2$, i.e., close to parallel to $\mathbf{e}_1$ !!

## ☆ Orthogonal Vectors

Def. • Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ are said to be orthogonal if $\mathbf{x}^T\mathbf{y} = 0$. So, the zero vector $\mathbf{0}$ is orthogonal to any vector.

• Two sets of vectors $X, Y$ are said to be orthogonal if $\forall \mathbf{x} \in X,\ \forall \mathbf{y} \in Y,\ \mathbf{x}^T\mathbf{y} = 0$.

• A set of vectors $S$ is said to be orthogonal if $\forall \mathbf{x} \in S,\ \forall \mathbf{y} \in S,\ \mathbf{x} \neq \mathbf{y}$ $\mathbf{x}^T\mathbf{y} = 0$.

- A set of vectors $S$ is said to be orthonormal if $S$ is orthogonal and $\forall x \in S$, $\|x\|_2 = 1$.

even more balanced!

**Thm** The vectors in an orthogonal set $S$ are linearly independent.

(Proof) Let $S = \{v_1, \cdots, v_n\}$
Suppose they are not lin. indep.
Then $\exists\, v_k \in S$ s.t. $v_k \neq \mathbb{0}$ and

$$v_k = \sum_{\substack{i=1 \\ i \neq k}}^{n} c_i v_i \quad \text{with } \mathbb{c} \neq \mathbb{0}$$

$$\mathbb{c} = [c_1, \cdots, c_{k-1}, c_{k+1}, \cdots, c_n]^T$$

Since $S$ is an orthogonal set,
$v_j^T v_i = 0$ for $\forall j \neq i$.

But $$v_k^T \Big( \sum_{\substack{i=1 \\ i \neq k}}^{n} c_i v_i \Big) = \sum_{\substack{i=1 \\ i \neq k}}^{n} c_i \underbrace{v_k^T v_i}_{=0} = 0$$

$\Leftrightarrow v_k^T v_k = 0$
$\Leftrightarrow \|v_k\|^2 = 0 \Leftrightarrow v_k = \mathbb{0}$ # contradiction!

## ☆ Components of a vector

SLOGAN " Inner products can be used to decompose arbitrary vectors into orthogonal components! "

Suppose $\{q_1, \cdots, q_n\} \subset \mathbb{R}^m$ is an orthonormal set. $q_j \in \mathbb{R}^m$, $1 \le j \le n$.

Let $v$ be an arbitrary vector in $\mathbb{R}^m$.

$$r = v - (q_1^T v) q_1 - (q_2^T v) q_2 - \cdots - (q_n^T v) q_n$$

$\uparrow$ residual vector is $\perp$ to $\{q_1, \cdots, q_n\}$

<u>why?</u>

$$q_j^T r = q_j^T v - (q_1^T v)\underbrace{q_j^T q_1}_{=0} - \cdots - (q_{j-1}^T v)\underbrace{q_j^T q_{j-1}}_{=0}$$
$$- (q_j^T v)\underbrace{q_j^T q_j}_{=1} - (q_{j+1}^T v)\underbrace{q_j^T q_{j+1}}_{=0} - \cdots - (q_n^T v)\underbrace{q_j^T q_n}_{=0}$$

$$= q_j^T v - q_j^T v = 0$$

This is true for any $j = 1, \cdots, n$

$$\Rightarrow \quad v = r + \sum_{i=1}^{n} (q_i^T v)\, q_i$$

any vector in $\mathbb{R}^m$

$$= r + \sum_{i=1}^{n} (q_i q_i^T)\, v$$

$\perp$

$$= r + QQ^T v$$

where $Q := [q_1 \ \cdots \ q_n] \in \mathbb{R}^{m \times n}$

If $\{q_1, \cdots, q_n\}$ is a basis of $\mathbb{R}^m$,
then $n = m$ and $r = 0$

i.e., $\displaystyle v = \sum_{i=1}^{m} (q_i^T v)\, q_i = \sum_{i=1}^{n} (q_i q_i^T)\, v$

In fact, $v = QQ^T v$, i.e.,

$$QQ^T = I$$

<u>Def.</u> A square matrix $Q \in \mathbb{R}^{m\times m}$ is said to be orthogonal if

$$Q^T = Q^{-1}$$

should be called orthonormal

i.e., $Q^TQ = QQ^T = I$

<u>Note</u>: If $Q = [q_1 \cdots q_n] \in \mathbb{R}^{m\times n}$ with $m > n$ and these vectors are orthonormal, then it is always true that $Q^TQ = I_{n\times n}$ but $QQ^T \neq I_{m\times m}$ unless $m = n$

e.g.,

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix} \text{ then } Q^TQ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2}$$

But, 

$$QQ^T = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 5/6 & 1/3 & -1/6 \\ 1/3 & 1/3 & 1/3 \\ -1/6 & 1/3 & 5/6 \end{bmatrix} \neq I_{3\times 3}$$

<u>Why?</u> $\Rightarrow$ Next lecture on <u>Orthogonal Projector</u>.

# ☆ Multiplication by an ortho. matrix

$$\begin{tikzcd}[column sep=huge] \begin{array}{c} y : \text{coef's} \\ \text{of expansion} \\ \text{in } \{e_1, \cdots, e_m\} \end{array} \arrow[r, shift left, "Q^T\cdot"] & \begin{array}{c} Q^T y : \text{coef's} \\ \text{of expansion of} \\ y \text{ in } \{q_1, \cdots, q_m\} \end{array} \arrow[l, shift left, "Q\cdot"] \end{tikzcd}$$

Note that $\| y \| = \| Q^T y \|$ !

i.e., isometry !

why?

$$\| Q^T y \|^2 = (Q^T y)^T (Q^T y)$$
$$= y^T \underbrace{Q\, Q^T}_{I} y$$
$$= y^T y = \| y \|^2 \; !!$$

Compare this with the general situation we discussed before : $A \in \mathbb{R}^{m \times m}$, nonsingular

$$\begin{tikzcd}[column sep=huge] \begin{array}{c} y : \text{coef's} \\ \text{of expansion} \\ \text{in } \{e_1, \cdots, e_m\} \end{array} \arrow[r, shift left, "A^{-1}\cdot"] & \begin{array}{c} A^{-1} y : \text{coef's} \\ \text{of expansion of} \\ y \text{ in } \{a_1, \cdots, a_m\} \end{array} \arrow[l, shift left, "A\cdot"] \end{tikzcd}$$