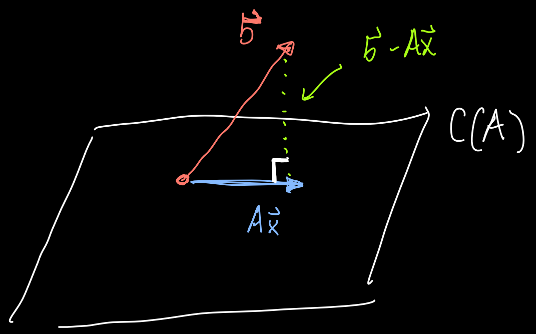
Last class... solving $A\vec{x} = \vec{b}$ when $A$ is invertible

"never invert a matrix"

- computational cost
- accuracy.

Solving $A\vec{x} = \vec{b}$ when there is no unique solution

If $A$ is $n \times p$, we can try to minimize $\|A\vec{x} - \vec{b}\|_2^2$




The projection of $\vec{b}$ into $C(A)$, $A\vec{x}$, will be orthogonal to $\vec{b} - A\vec{x}$

$$A^T(\vec{b} - A\vec{x}) = \vec{0} \Rightarrow A^T\vec{b} - A^TA\vec{x} = \vec{0}$$

$$\Rightarrow A^T\vec{b} = A^TA\vec{x} \quad \text{(normal equation)}$$

It turns out minimizing $\|A\vec{x} - \vec{b}\|_2^2$ is equivalent to solving the normal equation $A^TA\vec{x} = A^T\vec{b}$ for $\vec{x}$ when $A^TA$ is invertible.

$$\|A\vec{x}-\vec{b}\|_2^2 = (A\vec{x}-\vec{b})^T(A\vec{x}-\vec{b})$$

$$= x^TA^TA\vec{x} - 2\vec{b}^TA\vec{x} + \vec{b}^T\vec{b}$$

$X\vec{\beta} = \vec{Y}$

$\hat{\vec{\beta}} = (X^TX)^{-1}X^T\vec{Y}$

How do we solve normal equations? $A^TA\vec{x} = A^T\vec{b}$

When $A^TA$ is invertible, $\boxed{\hat{\vec{x}} = (A^TA)^{-1}A^T\vec{b}}$

## Cholesky Decomposition

theorem if $S \in \mathbb{R}^{n\times n}$ is symmetric positive definite, there exists a unique lower triangular matrix with positive diagonal entries $L \in \mathbb{R}^{n\times n}$ such that $S = LL^T$

def $L$ is called a Cholesky factor

$LL^T$ is the Cholesky decomposition

Q. When is $A^TA$ symmetric? Always

When is $A^TA$ symmetric pos. def?

when $A$ is full rank (linearly independent columns)

$A^T A = LL^T$

Solve $A^T A \vec{x} = A^T \vec{b}$

$LL^T \vec{x} = \vec{c}$ where $\vec{c} = A^T \vec{b}$

Solve $L\vec{y} = \vec{c}$ where $\vec{y} = L^T \vec{x}$ (forward)

Solve $L^T \vec{x} = \vec{y}$ (backward)

Suppose we want to simulate $\vec{X} \sim N(\vec{0}, \Sigma)$

Very easy to sample $\vec{Z} \sim N(\vec{0}, I_n)$

$L\vec{Z} \sim N(\vec{0}, LL^T)$

---

Solving using QR decomposition.

What if we are worried about stability?

The condition number of $A^T A$ is $\|A^T A\| \|{A^T A}^{-1}\| = \frac{\sigma_1^2}{\sigma_n^2}$

where $\sigma_1$ is max singular value of $A$.

Instead of solving $\vec{x} = (A^T A)^{-1} A^T \vec{b}$, we can use the QR decomposition:

A matrix can be decomposed $A = QR$

where $Q$ is orthogonal $\qquad K_2(Q) = 1$

$R$ is upper triangular

$$\hat{x} = (A^T A)^{-1} A^T \vec{b} = (R^T Q^T Q R)^{-1} R^T Q^T \vec{b}$$

$$= (R^T R)^{-1} R^T Q^T \vec{b}$$

$$= R^{-1} Q^T \vec{b}$$

The benefit here is accuracy.

---

What if $A^T A$ is not invertible?

If $A$ is invertible the solution to $A\vec{x} = \vec{b} = A^{-1}\vec{b}$

If $A$ is not square, we can still compute its

<u>pseudoinverse</u>

What would it mean to invert a non-square matrix?

Let $A^+$ be a pseudoinverse of $A$.

desired properties

- If $A$ is invertible $A^+ = A^{-1}$
- If $A$ is $m$ by $n$, $A^+$ is $n$ by $m$.
- $A^+ A\vec{x} = \vec{x}$ when $\vec{x}$ is in row space of $A$
- $AA^+\vec{b} = \vec{b}$ when $\vec{b}$ is in column space of $A$.

$AA^{-1} = I$

$A^{-1}A = I$

$\begin{bmatrix} A \end{bmatrix}$ $m \times n$

$m > n$

$\begin{bmatrix} A^+ \end{bmatrix}$ $n \times m$

$C(A) \subset \mathbb{R}^m$

$C(A^T) \subset \mathbb{R}^n$

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

$AA^+\vec{b} = \vec{b}$ when $\vec{b} \in C(A)$

$A^+A\vec{x} = \vec{x}$ when $\vec{x} \in C(A^T)$