Sect 6.1 (Omitted by accident)

Thm

Let A be an $m \times n$ matrix, then the orthogonal complement of the row space of A is the null space, and orthogonal complement of the column space is the null space of A^T:

$$(\text{Row } A)^{\perp} = \text{Nul } A \qquad (1)$$

and

$$(\text{Col } A)^{\perp} = \text{Nul } A^T \qquad (2)$$

Pf.

Let $\underline{x}$ be in Nul A, then we have $A\underline{x} = \underline{0}$. If we write $A\underline{x}$

$$A\underline{x} = \begin{vmatrix} \underline{a}_1 \\ \underline{a}_2 \\ \vdots \\ \underline{a}_m \end{vmatrix} \underline{x} = \begin{vmatrix} \underline{a}_1 \underline{x} \\ \underline{a}_2 \underline{x} \\ \vdots \\ \underline{a}_m \underline{x} \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}$$

$\underline{a}_i$ is ith row of A

Here $$\underline{a}_i \underline{x} = \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{vmatrix} = \underline{a}_i \cdot \underline{x} = 0$$

So $\underline{x}$ is orthogonal to the ith row of A for $i = 1 \cdots m$. Conversely, if we assume $\underline{x}$ is orthogonal to Row A, then $A\underline{x} = 0$ and

$\underline{x}$ is in Nul A. $\underline{x}$ orthogonal to all vectors in Row A $\Rightarrow$ $\underline{x}$ is in $(\text{Row } A)^{\perp}$
This shows (1).

To see (2), we note that its also true for A^T, so we have

$$(\text{Col } A)^{\perp} = (\text{Row } A^T)^{\perp} = \text{Nul } A^T$$

This proves (2).

Sect 6.4 (Cont...)

Recall: we introduced the Gram-Schmidt process as a way to produce an orthonormal basis for a subspace W.

We now consider this for the subspace $W = \text{Col } A$ for an $m \times n$ matrix A.

Thm

If A is an $m \times n$ matrix with lin. independent. columns, then A can be factored as $A = QR$ where Q is $m \times n$ and its columns form an orthonormal basis for Col A. R is an $n \times n$ upper triangular matrix with positive diagonal entries.

Pf

The columns of $A = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ form a basis for Col A. Using Gram-Schmidt, construct an orthonormal basis $\{\underline{u}_1 \cdots \underline{u}_n\}$ for Col A such that $\text{span}\{\underline{u}_1 \cdots \underline{u}_k\} = \text{span}\{\underline{a}_1 \cdots \underline{a}_k\}$ for all $k = 1 \cdots n$.

Set $Q = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix}$, an $m \times n$ matrix

Since $\underline{a}_k \in \text{span}\{\underline{a}_1 \cdots \underline{a}_k\} = \text{span}\{\underline{u}_1 \cdots \underline{u}_k\}$ there exist weights $r_{1k} \cdots r_{kk}$ such that

$$\underline{a}_k = r_{1k}\underline{u}_1 + r_{2k}\underline{u}_2 + \cdots + r_{kk}\underline{u}_k$$

Assume $r_{kk} > 0$ (if not, multiply $\underline{u}_k$ and r_{kk} by -1) and set the vector

$$\underline{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} 1 \cdots k \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} k+1 \cdots m \end{array}$$

Then $Q\underline{r}_k = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_n \end{bmatrix} \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$$= \underbrace{r_{1k}\underline{u}_1 + r_{2k}\underline{u}_2 + \cdots + r_{kk}\underline{u}_k}_{} + \underbrace{0 \cdot \underline{u}_{k+1} + \cdots + 0 \cdot \underline{u}_n}_{\underline{0}}$$

$$= \underline{a}_k$$

Let $R = \begin{vmatrix} \underline{r}_1 & \underline{r}_2 & \cdots & \underline{r}_n \end{vmatrix} = \begin{vmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ \vdots & 0 & r_{33} & \cdots & r_{3n} \\ & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{vmatrix}$ ← upper triangular

We get $QR = \begin{vmatrix} Q\underline{r}_1 & Q\underline{r}_2 & \cdots & Q\underline{r}_n \end{vmatrix}$

$= \begin{vmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \end{vmatrix} = A$

R is upper triangular and $r_{kk} > 0$ for $k = 1, \dots, n$.

↳ This means $\det(R) > 0 \Rightarrow R$ is invertible

Q is an $m \times n$ orthogonal matrix

Ex

Let $A = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \in \mathbb{R}^{4\times3}$. Compute $A = QR$.

Last class, we computed an orthonormal basis for $\{\underline{a}_1\ \underline{a}_2\ \underline{a}_3\}$, given by

$$\underline{u}_1 = \begin{vmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{vmatrix}, \quad \underline{u}_2 = \begin{vmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{vmatrix}, \quad \underline{u}_3 = \begin{vmatrix} 0 \\ -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{vmatrix}$$

So, we set $Q = \begin{vmatrix} \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \end{vmatrix} \in \mathbb{R}^{4\times3}$. To find R, we note the following

$$Q^T A = Q^T(QR) = (Q^T Q)R = R$$

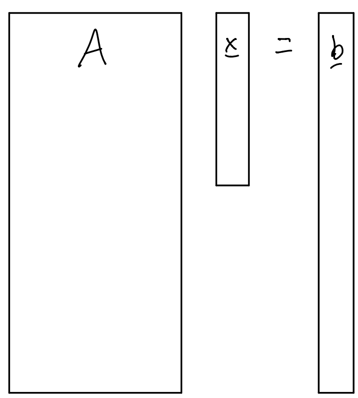
So we compute

$$R = \begin{vmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{vmatrix}$$

- See R is upp. tri. and $\det(R) = 2 \cdot (3/\sqrt{12}) \cdot (2/\sqrt{6}) > 0$ so R is invertible
- Check that $A = QR$ for yourself.

Sect 6.5: Least-Squares Problems

Consider a system of equations where A is $m \times n$, $\underline{x}$ is $n \times 1$, $\underline{b}$ is $m \times 1$, and $m >> n$



There are many more equations (rows) than variables (cols)
We call such system over-determined.

Oftentimes, over-determined systems have no exact $\underline{x}$ such that $A\underline{x} = \underline{b}$. Instead, we try to find the "closest" solution, i.e., find an $\underline{x}$ such that

$$A\underline{x} \approx \underline{b} \iff \underline{0} \approx \underline{b} - A\underline{x}$$

We define "closest" using a norm

$$\| \underline{b} - A\underline{x} \|$$

We want $\|b - A\underline{x}\|$ to be as small as possible. Note that

$$\| b - Ax \| = \left(\sum_{i=1}^{m} (b_i - [Ax]_i)^2 \right)^{1/2}$$

We call these problems least-squares problems because the solution is the least sum of squares.

<u>Def</u>

If A is $m \times n$ and $\underline{b}$ is in $\mathbb{R}^m$, a <u>least-squares</u> solution of $A\underline{x} = \underline{b}$ is $\hat{\underline{x}} \in \mathbb{R}^n$ such that

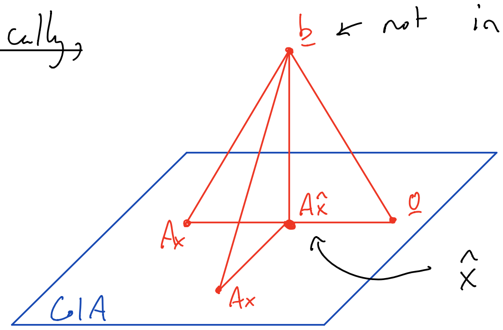
$$\| \underline{b} - A\hat{\underline{x}} \| \leq \| \underline{b} - A\underline{x} \|$$

for all $x \in \mathbb{R}^n$.

What does a least-squares solution look like?

Geometrically,

$\underline{b}$ ← not in Col A



$\hat{x}$ is the vector s.t. $A\hat{x}$ is closest to $\underline{b}$

$A\underline{x} \in \text{Col } A$ for all $\underline{x}$