Last time: Given linearly independent $\{v_1, \dots, v_m\}$ in $\mathbb{C}^n$,

define $u_k = \dfrac{\left(v_k - \sum_{i<k} \langle v_k, u_i \rangle u_i\right)}{\| v_k - \sum_{i<k} \langle v_k, u_i\rangle u_i \|_2}$ for $1 \le k \le m$

Then, $\{u_1, \dots, u_m\}$ is an o.n. set that has the same span as that of $\{v_1, \dots, v_m\}$. (o.n. = orthonormal)

Let $A_{m\times n} = [A_1 \cdots A_n]$

Gram-Schmidt algorithm: (B, C, T below can be thought of as $m\times n$ 0 matrices initially)

for $j = 1$ to n do

  for $i = 1$ to $j-1$ do

    $t_{ij} \leftarrow \langle A_j, B_i \rangle$

  end do

  $C_j \leftarrow A_j - \sum_{i<j} t_{ij} B_i$ (*)

  $t_{jj} \leftarrow \|C_j\|_2$

  $B_j \leftarrow t_{jj}^{-1} C_j$

end do

(for $j=1$, we have $i = 1$ to 0; this means do not enter the loop)

$i < j = 1 \Rightarrow$ skip sum.

$t_{ij} = (T)_{ij}$

(Using the notation from the above algorithm)

Theorem: If $A_{m\times n}$ is of rank n, then

$$A = BT$$

where B is an $m\times n$ matrix whose columns are o.n. and T is an $n\times n$ upper triangular matrix with positive diagonal.

Pf Columns of B are o.n. by previous remarks.

Moreover,

$$A_j = \sum_{i=1}^{j} \langle A_j, B_i \rangle B_i \qquad \text{since } \operatorname{span}\{A_1 \dots A_j\} = \operatorname{span}\{B_1 \dots B_j\} \text{ and } \{B_1, \dots B_j\} \text{ is o.n.}$$

$$= \sum_{i=1}^{j-1} t_{ij} B_i + \langle A_j, B_j \rangle B_j \qquad \left(\text{Recall: } x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i\right)$$

$$\langle A_j, B_j \rangle \overset{(*)}{=} \left\langle C_j + \sum_{i<j} t_{ij} B_i, B_j \right\rangle = \langle C_j, B_j \rangle \quad (\text{since } \langle B_i, B_j \rangle = \delta_{ij}, \text{ so for } i<j \ \langle B_i, B_j\rangle = 0)$$

$$= \langle C_j, C_j \rangle / t_{jj} = t_{jj}^2 / t_{jj} = t_{jj}$$

Thus, $A_j = \sum_{i=1}^{j} t_{ij} B_i \iff A = BT \iff A_{ij} = \sum_k B_{ik} T_{kj}$

(Exercise)

Least-Squares Problems

Let $A_{m\times n}$ and consider the system $Ax = b$ where $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$.

Suppose $\operatorname{rank}(A) = n$. Thus, $n \le m$.

In general $Ax = b$ may not have a solution.

In fact,

$$Ax = b \text{ has a solution} \iff b \in \operatorname{colspace}(A)$$

$Ax = b$

$[A_1 \cdots A_n]\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = b$

$A_1x_1 + A_2x_2 + \dots + A_nx_n = b$ — lin. comb. of vectors $A_1, \dots A_n$ which are column vectors of A

If $b \notin \operatorname{colspace}(A)$, then we might want x s.t. $\|b - Ax\|_2$ is minimized.

This is the least-squares "solution" of $Ax = b$.

$\|b - Ax\|_2$ is minimized $\iff \|b-Ax\|_2^2$

Lemma If x satisfies $A^*(Ax - b) = 0$ then x solves the least squares problem.

$A_{m\times n}$, $(A)_{n\times 1}$, $(A)^*_{1\times m}$

Pf Note that $A^*(Ax-b) = [A_1^*(Ax-b)_{m\times 1} \ A_2^*(Ax-b) \cdots A_n^*(Ax-b)] = 0$

$\Rightarrow Ax - b \perp \operatorname{colspace}$ of A. $\langle x, y \rangle = y^* x$, $A_i^*(Ax-b) = \langle Ax-b, A_i \rangle = 0$

A_i: i^{th} column

Let $y \in \mathbb{C}^n$. Then $A(x-y) \in \operatorname{colspace} A$

Thus, $\langle b - Ax, A(x-y) \rangle = 0$

$$\Rightarrow \|b - Ay\|_2^2 = \|b - Ax + Ax - Ay\|_2^2$$

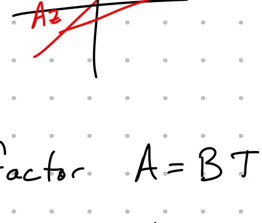
$$= \|b - Ax\|_2^2 + \|A(x-y)\|_2^2 \quad \text{why?}$$

$$\ge \|b - Ax\|_2^2$$

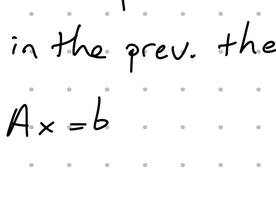
Pythagorean theorem

$\|a+b\|^2 = \|a\|^2 + \|b\|^2$, $a \perp b$

Picture



$b - Ax \perp \operatorname{colspace}(A)$



$A = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$

Claim: If we factor $A = BT$ as in the prev. theorem, then the LS solution of $Ax = b$ is the exact solution of

$$Tx = (B^*B)^{-1} B^* b$$

Recall T is $n\times n$

Pf $A^*Ax = (BT)^*(BT)x = T^*B^*BTx$

$= T^*B^*B(B^*B)^{-1}B^*b$

$= T^*B^*BB^{-1}B^{*-1}B^*b = T^*B^*b = A^*b$

so $A^*Ax - A^*b = 0$

$\Rightarrow A^*(Ax - b) = 0 \Rightarrow x$ is the LS solution to $Ax = b$.

5.4 Singular Value Decomposition and Pseudoinverses

Recall A is symmetric if $A^T = A$
skew-sym. if $A^T = -A$. (for real matrices)

Def^n A is Hermitian if $A^* = A$
skew-Hermitian if $A^* = -A$ (for complex matrices)

Recall $A \in \mathbb{R}^{n\times n}$ is pos. def if $x^T A x > 0$ $\forall x \neq 0 \in \mathbb{R}^n$
pos. semidefinite if $x^T A x \geq 0$ $\forall x \in \mathbb{R}^n$

Def^n $A \in \mathbb{C}^{n\times n}$ is pos. def if $x^* A x > 0$ $\forall x \neq 0 \in \mathbb{C}^n$
is pos. semidefinite if $x^* A x \geq 0$ $\forall x \in \mathbb{C}^n$

Thm (SVD)
let $A \in \mathbb{C}^{m\times n}$. Then $\exists P \in \mathbb{C}^{m\times m}$, $D \in \mathbb{C}^{m\times n}$, $Q \in \mathbb{C}^{n\times n}$
where P and Q are unitary (i.e. $P^*P = I$ and $Q^*Q = I$)
and D is diagonal and $A = PDQ$.

Pf The matrix A^*A is $n\times n$ Hermitian since $(A^*A)^* = A^*A$

Also, $x^*(A^*A)x = (x^*A^*)(Ax) = (Ax)^*(Ax) = \langle Ax, Ax\rangle \geq 0$

That is A^*A is positive semidefinite.

(If B is positive semidefinite, eigenvalues of B are nonnegative)

Let $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ be the eigenvalues.

We may assume $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_r^2 \geq \sigma_{r+1}^2 = 0 = \dots = \sigma_n^2$

Let $\{u_1, \dots, u_n\}$ be o.n. set of eigenvectors for A^*A s.t.

$$A^*A\, u_i = \sigma_i^2 u_i$$
$$B\ x = \lambda\ x$$

Then
$$\|Au_i\|_2^2 = \langle Au_i, Au_i\rangle = (Au_i)^* Au_i = u_i^* A^* A u_i = \sigma_i^2 u_i^* u_i = \sigma_i^2$$

$\Rightarrow Au_i = 0$ for $i > r$. $Q^*Q = I \Rightarrow Q^{-1} = Q^*$

Define $Q = \begin{pmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_n^* \end{pmatrix}_{n\times n}$ $v_i = \sigma_i^{-1} A u_i$ $(1 \leq i \leq r)$

Then
$$v_i^* v_j = \sigma_j^{-1}\sigma_i^{-1} u_i^* A^* A u_j = \sigma_j^{-1}\sigma_i^{-1} u_i^* \sigma_j^2 u_j$$
$$= \sigma_j \sigma_i^{-1} u_i^* u_j = \sigma_j \sigma_i^{-1} \delta_{ij} = \delta_{ij}$$

Extend $\{v_1, \dots, v_r\}$ to an o.n. basis $\{v_1, \dots, v_r, v_{r+1}, \dots, v_m\}$
for $\mathbb{C}^m$.

Set $P = [v_1 \cdots v_m]$. $\Rightarrow P^*P = I \Rightarrow P^{-1} = P^*$

let $D \in \mathbb{C}^{m\times n}$ be defined by $D_{ii} = \sigma_i$ for $1 \leq i \leq r$
and 0 everywhere else.

$$A = PDQ$$
$$P^{-1}AQ^{-1} = D$$
$$P^*AQ^* = D$$

Then $(P^*AQ^*)_{ij} = v_i^* A u_j$

if $j > r$, Then $Au_j = 0$ so $(P^*AQ^*)_{ij} = 0$

if $j \leq r$, $v_i^* A u_j = \sigma_i^{-1} u_i^* A^* A u_j = \sigma_j^2 \sigma_i^{-1} u_i^* u_j = \sigma_i \delta_{ij}$

Thus, $P^*AQ^* = D$. Since P and Q are unitary, ($P^*P = I$; $Q^*Q = I$)
$A = PDQ$.

The numbers $\sigma_1, \dots, \sigma_n$ are called the singular values of A.