MATH 312
LECTURE 18

: Singular Value Decomposition.

I.) Recall that any (square real symmetric) positive semidefinite matrix A can be factorised as follows:

$A = QDQ^T$, where Q is orthogonal

D is diagonal with nonnegative entries.

This can be rewritten as

$A = \lambda_1 \vec{q}_1 \vec{q}_1^T + \dots + \lambda_r \vec{q}_r \vec{q}_r^T$, where

- $\lambda_1, \dots, \lambda_r \equiv$ nonzero eigenvalues of A (may be repeated)
- $\vec{q}_1, \dots, \vec{q}_r \equiv$ eigenvectors of A corresponding to $\lambda_1, \dots, \lambda_r$.
  (columns of Q: so orthonormal vectors).
- $r = \text{rank}(A)$ (The spectral theorem ensures that we can find r orthonormal eigenvectors).

<u>Remark</u>: In Q there are $n-r$ additional orthonormal columns, corresponding to $\lambda = 0$ (nullspace of A).

<u>Remark</u>: The order of the λ's in D has to be the same as the order of $\vec{q}$'s in Q.

→ Notice that we have decomposed A as a sum of r matrices with rank 1.

We could order these using the weights $\underline{\underline{\lambda}}$'s.

→ Each piece is a projection matrix onto the line given by the corresponding eigenvector.

→ Since A is symmetric, $\{\vec{q}_1, \dots, \vec{q}_r\}$ is both an orthonormal basis of the column space and row space.

$(C(A) = C(A^T)$ since $A = A^T$!$)$.

For same reason, $\{\vec{q}_{r+1}, \dots, \vec{q}_n\}$ is a basis for nullspace and left nullspace.

• Now, we want to generalise this decomposition for any real matrix A $m \times n$.

↳ Singular Value Decomposition.

<u>Remark</u>: We will be able to check that for (symmetric) positive semidefinite matrix the SVD is the same as the above one.

So now, we want to show that any real $m \times n$ matrix A of rank r can be decomposed as follows:

$A = U \Sigma V^T$, where

1) Σ is an $m \times n$ (pseudo) diagonal matrix, with nonnegative entries.

2) V is an $n \times n$ orthogonal matrix. (*)

3) U is a $m \times m$ orthogonal matrix

Let's first see <u>one way</u> to find such Σ, U, V that <u>always</u> works. Then we will prove why.

(*) <u>Reminder</u>: Orthogonal matrix means <u>orthonormal</u> cols. and rows.

• Recipe for $A = U\Sigma V^T$.

1) $\underline{\underline{\Sigma}}$

1.1) Find eigenvalues of A^TA or AA^T

<u>Remark</u>: The nonzero ones are the same for A^TA and AA^T.

[So use the smaller matrix!]

1.2) Order the nonzero eigenvalues:

$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$

1.3) Let $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, \dots, r$. $\|\sigma$'s $\equiv$ Singular Values of A.$\|$

1.4) Finally, construct Σ of size <u>$m \times n$</u>:

$$\Sigma = \begin{array}{c} \left[\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} r \\ \left.\vphantom{\begin{array}{c}a\\a\\a\end{array}}\right\} m-r \end{array} \\ \underbrace{\hphantom{\sigma_1\ \ddots\ \sigma_r}}_{r} \quad \underbrace{\hphantom{0\ \ddots\ 0}}_{n-r} \end{array}$$

Example:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (5 \times 4)$$

Singular values of A? Σ?

$A^T A \to 4 \times 4$

$A A^T \to 5 \times 5$

$$\to A^T A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix} \to \lambda_1 = 9,\ \lambda_2 = 4,\ \lambda_3 = 1,\ \lambda_4 = 0.$$

Thus, $\sigma_1 = 3$, $\sigma_2 = 2$, $\sigma_1 = 1$ and

$$\Sigma = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

• Remark: Recall that rank(A) is equal to the number of singular values (three in this case).

2) $\underline{\underline{V}}$

2.1) Find r orthonormal eigenvectors of A^TA:

$$\begin{array}{ccccccc} \lambda_1 & \geq & \lambda_2 & \geq \dots \geq & \lambda_r & > 0 \\ | & & | & \dots & | & \\ \vec{v}_1 & & \vec{v}_2 & & \vec{v}_r & \end{array}$$

2.2) Find n−r orthonormal eigenvectors for $\lambda=0$ in A^TA..

<u>Remark</u>: • In both cases, Gram-Schmidt is needed if an eigenvalue is repeated.

• Vectors have to be unit length.

$$V = \left[\begin{array}{ccc|ccc} \vec{v}_1 & \dots & \vec{v}_r & \vec{v}_{r+1} & \dots & \vec{v}_n \end{array}\right] \text{ orthogonal } n \times n$$

$$\underbrace{\qquad\qquad}_{r} \quad \underbrace{\qquad\qquad}_{n-r}$$

→ The first r cols. have to be in order!!.

<u>Example</u>:

$$\lambda_1 = 9 \rightarrow A^TA - 9I = \begin{bmatrix} -9 & 0 & 0 & 0 \\ 0 & -8 & 0 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \dots \lambda_2, \lambda_3, \lambda_4$$

$$\hookrightarrow \quad V = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \end{array}$$

3) $\underline{\underline{U}}$

3.1) For $i = 1, \ldots, r$, let $\vec{u}_i = \dfrac{A\vec{v}_i}{\sigma_i}$.

3.2) Find $m-r$ orthonormal eigenvectors of AA^T corresponding to $\lambda = 0$:

$\vec{u}_{r+1}, \ldots, \vec{u}_m$ (i.e., find an orthonormal basis for $N(AA^T)$).

$$U = \left[\begin{array}{ccc|ccc} \dfrac{A\vec{v}_1}{\sigma_1} & \cdots & \dfrac{A\vec{v}_r}{\sigma_r} & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{array}\right] \quad \text{orthogonal } m \times m$$

(order matters!

these will be orthonormal (if not, previous mistake!).

Ex:

$$A\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \left(\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}\right)$$

$$\Rightarrow \vec{u}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad \text{We need } \vec{u}_4, \vec{u}_5 \in N(AA^T):$$

$$AA^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \vec{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \rightarrow U = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

[$\rightarrow$ Check: $A = U\Sigma V^T$?]

• We want to see now that such U, Σ, V always exist and that $A = U \Sigma V^T$.

1) $\underline{\underline{\Sigma}}$

1.1) $\text{rank}(A) = \text{rank}(AA^T) = \text{rank}(A^TA)$ →

nonzero eigenvalues of $\left| \begin{array}{l} A^TA \\ AA^T \end{array} \right.$ are the same (thus, with same multiplicity)

1.2) A^TA, AA^T positive semidefinite $\Rightarrow \lambda \geq 0$.

Therefore, there are r positive eigenvalues $\lambda_1, \ldots, \lambda_r$ ~~and they are~~ (maybe repeated),

so we can define the singular values $\sigma_i = \sqrt{\lambda_i}$ $(i = 1, \ldots, r)$.

2) $\underline{\underline{V}}$

A^TA is symmetric $\Rightarrow$ There is an orthonormal basis of eigenvectors.

So V would be orthogonal by construction.

3) $\underline{\underline{U}}$

$$U = \left[\frac{A\vec{v}_1}{\sigma_1} \cdots \frac{A\vec{v}_r}{\sigma_r} \;\vdots\; \underbrace{\vec{u}_{r+1} \cdots \vec{u}_m}_{} \right]$$

m-r orthonormal vectors in $N(AA^T)$. $\equiv$ m-r orthonormal eigenvectors of AA^T for $\lambda = 0$.

(possible since $\dim N(AA^T) = m - r$).

Need to check:

3.1) Are the first r columns of U orthonormal?

3.2) Are the first r orthogonal to the other m-r?

3.1) Let's check if: $(i, j \in \{1, \dots, r\})$

$$\underbrace{\frac{A\vec{v}_i}{\sigma_i}}_{\vec{u}_i} \cdot \underbrace{\frac{A\vec{v}_j}{\sigma_j}}_{\vec{u}_j} = \frac{(A\vec{v}_i)^T}{\sigma_i} \frac{A\vec{v}_j}{\sigma_j} = \frac{\vec{v}_i^T A^T A \vec{v}_j}{\sigma_i \sigma_j} = \frac{\vec{v}_i^T \lambda_j \vec{v}_j}{\sigma_i \sigma_j} =$$

$\vec{v}_j$ is eigenvector of A^TA by our construction of V

$$= \frac{\lambda_j}{\sigma_i \sigma_j} \vec{v}_i^T \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ \dfrac{\lambda_i}{\sigma_i^2} = 1 & \text{if } i = j \end{cases}$$

$\equiv$ $\{\vec{u}_1, \dots, \vec{u}_r\}$ are orthonormal vectors.

by our definition of σ_i.

3.2) Notice that

$$\begin{cases} \vec{u}_{r+1}, \dots, \vec{u}_m \in N(AA^T) = N(A^T) & \left(\begin{array}{c}\text{proved in previous} \\ \text{lectures}\end{array}\right) \\ \underbrace{\frac{A\vec{v}_1}{\sigma_1}}_{\vec{u}_1}, \dots, \underbrace{\frac{A\vec{v}_r}{\sigma_r}}_{\vec{u}_r} \in C(A) & \end{cases}$$

[Recall that $A\vec{x}$ is $x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n$, i.e., lin. comb. of the cols. of A.]

So, $\left.\begin{array}{l} \vec{u}_{r+1}, \dots, \vec{u}_m \in N(A^T) \\ \vec{u}_1, \dots, \vec{u}_r \in C(A) \end{array}\right\}$ and $C(A) = N(A^T)^{\perp}$,

so all the $\vec{u}_1, \dots, \vec{u}_r$ are perpendicular to all $\vec{u}_{r+1}, \dots, \vec{u}_m$.

- Finally, we have to check that $A = U\Sigma V^T$ holds:

V orthogonal $\Rightarrow V^T = V^{-1} \Rightarrow \{ A = U\Sigma V^T \Leftrightarrow AV = U\Sigma \}$.

So let's check $AV = U\Sigma$:

$$A \left[\vec{v}_1 \dots \vec{v}_r \,\middle|\, \underbrace{\vec{v}_{r+1} \dots \vec{v}_n}_{\substack{\in N(A^TA) \\ = N(A)}} \right] = \left[\vec{u}_1 \dots \vec{u}_r \,\middle|\, \vec{u}_{r+1} \dots \vec{u}_m \right] \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \underbrace{}_{n-r} \end{bmatrix} \}\, m-r$$

(proved in previous lecture; and was in the exam)

$\Downarrow$

$$\to \quad A\left[\vec{v}_1 \dots \vec{v}_r \,\middle|\, \vec{v}_{r+1} \dots \vec{v}_n\right] = \left[A\vec{v}_1 \dots A\vec{v}_r \,\middle|\, \underbrace{\vec{0} \dots \vec{0}}_{n-r}\right]$$

$$\to \quad \left[\vec{u}_1 \dots \vec{u}_r \,\middle|\, \vec{u}_{r+1} \dots \vec{u}_m\right]\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ \hline & & & \underbrace{\quad}_{n-r} \end{array}\right] = \left[\vec{u}_1 \dots \vec{u}_r \,\middle|\, \underbrace{\vec{0} \dots \vec{0}}_{n-r}\right]$$

Both things are equal if $\sigma_i \vec{u}_i = A\vec{v}_i$ for $i = 1, \dots, r$, but this is exactly our choice for the $\vec{u}$'s //.

- <u>Some remarks</u>

1) $\{\vec{v}_1, \dots, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_n\}$ are eigenvectors of A^TA corresponding to eigenvalues $\sigma_1^2, \dots, \sigma_r^2$ and 0 ($n-r$ times).

Indeed:

$$A^TA = V\Sigma^T U^T U \Sigma V^T = V \underbrace{\Sigma^T \Sigma}_{\substack{\text{diagonal} \\ n \times n}} V^T$$

$$\text{diagonal}_{n\times n} \to \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix} \Big\} n-r$$

This is nothing more than the spectral decomposition of A^TA (V plays the role of the Q)

2) Similarly, $\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}_{r+1}, \dots, \vec{u}_m\}$ is a basis of eigenvectors of AA^T corresponding to eigenvalues $\sigma_1^2, \dots, \sigma_r^2$ and 0 ($m-r$ times)

$$\hookrightarrow AA^T = U \Sigma V^T V \Sigma^T U^T = U \underbrace{\Sigma \Sigma^T}_{} U^T$$

Spectral decomposition of AAT

diagonal $m \times m$

$$\begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \Big\} m-r$$

$m-r$

3) We have a complete analogy with the spectral decomposition of positive semidefinite matrices:

$$A = U \Sigma V^T = \sigma_1 \vec{u}_1 \vec{v}_1^T + \dots + \sigma_r \vec{u}_r \vec{v}_r^T$$

rank 1 matrices

4)

$$V = \Bigg[\underbrace{\vec{v}_1 \dots \vec{v}_r}_{\substack{\text{Basis for} \\ \text{row space} \\ C(A^T)}} \Bigg| \underbrace{\vec{v}_{r+1} \dots \vec{v}_n}_{\substack{\text{Basis for} \\ \text{nullspace} \\ N(A)}} \Bigg] \qquad U = \Bigg[\underbrace{\vec{u}_1 \dots \vec{u}_r}_{\substack{\text{Basis for} \\ \text{column space} \\ C(A)}} \Bigg| \underbrace{\vec{u}_{r+1} \dots \vec{u}_m}_{\substack{\text{Basis for} \\ \text{left-nullspace} \\ N(A^T)}} \Bigg]$$

II.) The Pseudoinverse, A^+

We are going to use the SVD of A to solve two previous problems of the course:

1) If $A\vec{x} = \vec{b}$ has infinite solutions, which one is the shortest length solution?

2) If A has linearly dependent columns, how to solve the least-squares problem associated to $A\vec{x} = \vec{b}$?

• <u>Def</u>: Pseudoinverse of A

$$A^+ = V\Sigma^+ U^T, \text{ where } \Sigma^+ = \begin{array}{cc} \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & \end{bmatrix} & \}\, n-r \\ \underbrace{}_{m-r} & \end{array}.$$

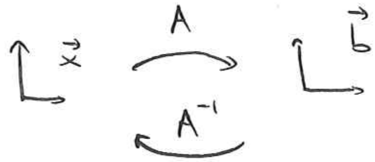
(Notice Σ^+ is $\underline{n \times m}$).

• <u>Remark</u>: If A is invertible, $A^+ = A^{-1}$.

↳ A invertible $\Rightarrow n = r \Rightarrow U, \Sigma, V$ square $n \times n$, and

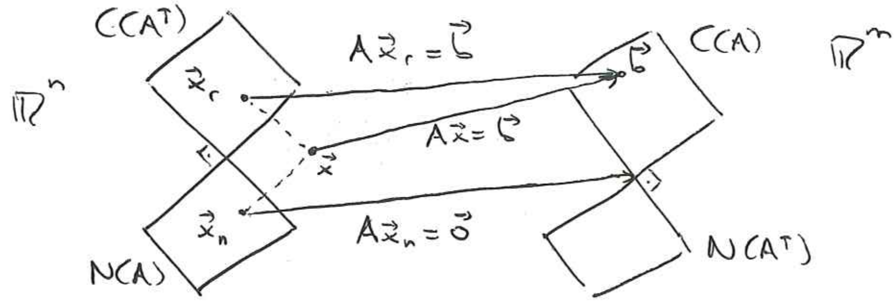
$$A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T \equiv V\Sigma^+ U^T = A^+$$

One of the easiest ways to think about the inverse, A^{-1}, is as the solution of $A\vec{x}=\vec{b}$.



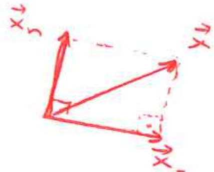
That is, knowing $\vec{b}$, we can find $\vec{x}$ as $A^{-1}\vec{b}$.

This is not possible if $A\vec{x}=\vec{b}$ has infinite solutions:



When $A\vec{x}=\vec{b}$ has infinite solutions, all the solutions are of the form $\vec{x}=\vec{x}_p+\vec{x}_n$, that is, we pick any solution and we can modify it by adding something in the nullspace.

How to decide which one to pick? It seems natural to choose the one with shortest length. That is the one (and only one) "living" in the row space, $\vec{x}_r$.



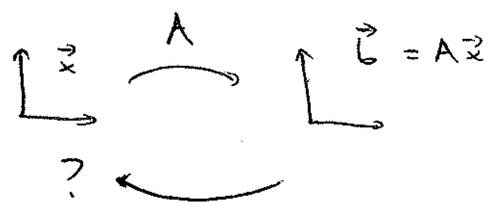
$\|\vec{x}\|^2=\|\vec{x}_r\|^2+\|\vec{x}_n\|^2$

$$\left.\begin{array}{l}\vec{x}=\vec{x}_r+\vec{x}_n \\ \text{with} \\ \underline{\underline{\vec{x}_r\cdot\vec{x}_r=0}}\end{array}\right\} \Rightarrow \|\vec{x}\|^2=\|\vec{x}_r\|^2+\underbrace{\|\vec{x}_n\|^2}_{\text{choose this to be } \vec{0}\,!}$$

In all this, it is crucial the fact that $C(A^T)$ and $N(A)$ are orthogonal complements.

Since both spaces are orthogonal to each other, given $\vec{x} \in \mathbb{R}^n$, we can find $\vec{x}_r$ as the projection of $\vec{x}$ onto $C(A^T)$ (since $N(A) \perp C(A^T)$, the part of $\vec{x} \in N(A)$ is "killed" by this projection).

- In summary, it is natural to define the inverse of the process in the picture,



by choosing $\vec{x}_r$.

- In other words, we want the solution of $A\vec{x} = \vec{b}$ to be $\vec{x}_r$, when this system has infinite solutions.

We are now going to see that $\boxed{\vec{x}_r = A^+ \vec{b}}$

• Is $A^+\vec{b} = \vec{x}_r$?

That is, is $A^+\vec{b}$ the projection of $\vec{x}$ onto $C(A^T)$?

Let's see if:

$$A\vec{x} = A\vec{b} \rightarrow A^+A\vec{x} = A^+\vec{b}, \text{ so, is } A^+A\vec{x} = \text{proj}_{C(A^T)}\vec{x} \text{ ?}$$

• $A^+A = V\Sigma^+U^TU\Sigma V^T = V\Sigma^+\Sigma V^T,$

$$\underset{(n\times m)}{\Sigma^+}\underset{(m\times n)}{\Sigma} = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & 1/\sigma_r \\ & & \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \sigma_r \\ & & \end{bmatrix} = \begin{array}{c} \left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & \end{array}\right] \begin{matrix} \\ \\ \\ \}n-r \end{matrix} \\ \hspace{3em}\underbrace{\hspace{2em}}_{n-r} \end{array}$$

$$V\Sigma^+\Sigma V^T = \left[\begin{array}{ccc|ccc} \vec{v}_1 & \cdots & \vec{v}_r & \vec{v}_{r+1} & \cdots & \vec{v}_n \end{array}\right]\left[\begin{array}{ccc|c} 1 & & & \\ & \ddots & & \\ & & 1 & \\ \hline & & & \end{array}\right]\left[\begin{array}{c} \vec{v}_1^T \\ \vdots \\ \vec{v}_r^T \\ \hline \vec{v}_{r+1}^T \\ \vdots \\ \vec{v}_n^T \end{array}\right] =$$

$$= \vec{v}_1\vec{v}_1^T + \dots + \vec{v}_r\vec{v}_r^T \equiv \text{Projection onto } \{\vec{v}_1, \dots, \vec{v}_r\}$$

the space spanned by

But recall that

$\{\vec{v}_1, \dots, \vec{v}_r\}$ was an orthonormal basis for $C(A^T)$!

• In summary, for $A\vec{x} = \vec{b}$, the solution $A^+\vec{b}$ gives

→ Shortest length solution.

→ Only solution in $C(A^T)$.

II.2) Least Squares.

Recall that if $A\vec{x} = \vec{b}$ doesn't have a solution (that is, $\vec{b}$ is not in the column space), then we solve instead

$$\left\| A\vec{x} = \text{proj}_{C(A)}\vec{b} \right\|.$$

We saw in chapter 4. that one way of solving this <u>new</u> system is by the normal equations:

$$A^TA\vec{x} = A^T\vec{b} \underset{\substack{\uparrow \\ \text{If } A^TA \text{ invertible}}}{\Longrightarrow} \vec{x} = (A^TA)^{-1}A^T\vec{b} \quad \left(\begin{array}{l}\text{indeed,} \\ A\vec{x} = \underbrace{A(A^TA)^{-1}A^T\vec{b}}_{\text{proj}_{C(A)}\vec{b}}\end{array}\right)$$

What if A^TA is not invertible? That is, what if the system has infinite solutions? $\left(\begin{array}{l}\text{always has at least one, as} \\ \text{proj}_{C(A)}\vec{b} \in C(A) \text{ obviously}\end{array}\right)$.

$\left[A^TA \text{ is not invertible if cols. of } A \text{ are lin. dependent}\right]$.

Well, now we know that for systems with infinite solutions the pseudoinverse is concerned.

Let's see that $\left\| \vec{x} = A^+ \vec{b} \right.$ is also the least-squares solution $\left. \right\|$ (with shortest length if there are many).

$\Rightarrow$ If $A\vec{x} = \vec{b}$ no solution $\to$ Solve instead $A\vec{x} = \text{proj}_{C(A)} \vec{b}$.

Proceeding as in -197-, one can find that

$$AA^+ = U \Sigma \Sigma^+ U^T = \vec{u}_1 \vec{u}_1^T + \dots + \vec{u}_r \vec{u}_r^T \equiv$$ Projection matrix onto subspace spanned by $\{\vec{u}_1, \dots, \vec{u}_r\}$

That is,

$$\left\| AA^+ \equiv \text{projection matrix onto } C(A) \right\|$$

So we want to solve $A\vec{x} = AA^+ \vec{b}$.

$\Rightarrow$ Solving $A\vec{x} = AA^+ \vec{b} \implies \vec{x} = A^+ A A^+ \vec{b}$

<u>Exercise</u>: Check that $A^+ = A^+ A A^+$.

Also check that $A = AA^+A$.

• <u>Summary</u>: For $A\vec{x}=\vec{b}$, the "solution" $\vec{x}=A^{+}\vec{b}$ gives:

1) Exact solution if $A\vec{x}=\vec{b}$ is solvable.

   1.1) If solution is unique, $A^{+}\vec{b}=A^{-1}\vec{b}$.

   1.2) If infinite solutions, $A^{+}\vec{b}$ is the shortest one.
   (<u>the</u> solution in $C(A^{T})$) ~~is the only solution~~

2) Least square solution to $A\vec{x}=\vec{b}$ if it is not solvable.

   2.1) If cols. of A are l.i., then $A^{+}\vec{b}=(A^{T}A)^{-1}A^{T}\vec{b}$.

   2.2) If cols. of A are l.i, then $A^{+}\vec{b}$ is the shortest least-squares solution.