A^+ = the pseudo-inverse of A

$\boxed{\text{called the Moore-Penrose generalized inverse}}$

First introduce notation:

Partitioned vectors:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_m \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} y_1 \\ \vdots \\ y_r \end{matrix}}\right\} \tilde{y} \\ \left.\vphantom{\begin{matrix} y_{r+1} \\ \vdots \\ y_m \end{matrix}}\right\} \bar{y} \end{matrix} = \begin{pmatrix} \tilde{y} \\ \bar{y} \end{pmatrix}$$

$$\|y\|^2 = \sum_{i=1}^{r} y_i^2 + \sum_{i=r+1}^{m} y_i^2 = \|\tilde{y}\|^2 + \|\bar{y}\|^2$$

Partitioned matrices / vectors:

$$Ax = \left[\begin{array}{c|c} B & C \\ \hline D & E \end{array}\right] \begin{pmatrix} \tilde{x} \\ \hline \bar{x} \end{pmatrix} = \begin{pmatrix} B\tilde{x} + C\bar{x} \\ \hline D\tilde{x} + E\bar{x} \end{pmatrix}$$

Derivation of A^+

$$\xrightarrow{x} \boxed{Ax = y} \longrightarrow y_d$$

Ideally: choose x so $Ax = y_d$

$\boxed{\begin{array}{l} \text{possible if } y_d \in \mathcal{R}[A] \\ \text{impossible if } y_d \notin \mathcal{R}[A] \end{array}}$

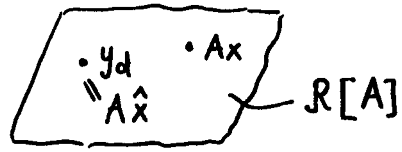
Suggests:

choose x to min $\|y_d - Ax\|$

$\hat{x}$

If $y_d \in \mathcal{R}[A]$:

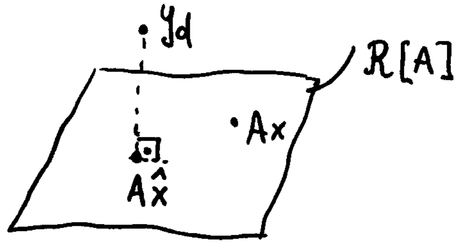
$\|y_d - A\hat{x}\| = 0$



If $y_d \notin \mathcal{R}[A]$:

$\|y_d - A\hat{x}\| > 0$

$\boxed{\text{but as small as possible}}$



Computation of an $\hat{x}$ aided by:

(7.0) **Th** For $0 \neq A \in \mathbb{R}^{m \times n}$, $\exists \perp P, \perp Q$

so:

$$A = P \left[\begin{array}{c|c} \hat{A}_{r\times r} & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right] Q^T$$

where: $\hat{A}_{r\times r} \in \mathbb{R}^{r\times r}$ is invertible

$P \in \mathbb{R}^{m\times m}$

$Q \in \mathbb{R}^{n\times n}$

and $0_{m\times n}$ denotes an $m\times n$ matrix full of zeros

---

Also minimizing $\|y_d - Ax\|$

equivalent to

minimizing $\|y_d - Ax\|^2$

This is the general case: one might actually get:

$A = P\begin{bmatrix}\hat{A}\\ 0\end{bmatrix}Q^T$, or $A = P[\hat{A}\ 0]Q^T$, or $A = P\hat{A}Q^T$

depending on A.

$$\boxed{\underset{x\in\mathbb{R}^n}{\text{Minimize}}\ \|y_d - Ax\|^2}$$

$\|y_d - Ax\|^2$

$\| \ \boxed{\|z\|^2 = \|P^Tz\|^2 \text{ as } P \text{ is } \perp}$

$\|P^T[y_d - Ax]\|^2$ with $A = P\begin{bmatrix}\hat{A} & 0\\ 0 & 0\end{bmatrix}Q^T$

$$= \|\underbrace{P^Ty_d}_{\begin{bmatrix}\tilde{y}_d\\ \bar{y}_d\end{bmatrix}} - \begin{bmatrix}\hat{A} & 0\\ 0 & A\end{bmatrix}\underbrace{Q^Tx}_{\begin{bmatrix}\widetilde{Q^Tx}\\ \overline{Q^Tx}\end{bmatrix}}\|^2$$

$\widetilde{Q^Tx}$ just denotes the top part of the partitioned vector Q^Tx, etc.

$$= \left\|\begin{bmatrix}\tilde{y}_d\\ \bar{y}_d\end{bmatrix} - \begin{bmatrix}\hat{A} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\widetilde{Q^Tx}\\ \overline{Q^Tx}\end{bmatrix}\right\|^2$$

$$= \left\|\begin{bmatrix}\tilde{y}_d - \hat{A}(\widetilde{Q^Tx})\\ \bar{y}_d\end{bmatrix}\right\|^2$$

$$= \underbrace{\|\tilde{y}_d - \hat{A}(\widetilde{Q^Tx})\|^2}_{\geq 0} + \|\bar{y}_d\|^2 \geq \|\bar{y}_d\|^2,\ \forall x$$

$$= \|\bar{y}_d\|^2 \text{ whenever } \tilde{y}_d = \hat{A}(\widetilde{Q^Tx})$$

Hence

$$\min_{x \in \mathbb{R}^n} \| y_d - Ax \|^2 = \| \bar{y}_d \|^2$$

$$\arg\min_{x \in \mathbb{R}^n} \| y_d - Ax \|^2 = \{ x \in \mathbb{R}^n : \tilde{y}_d = \hat{A} (\widetilde{Q^T x}) \}$$

$$= \{ x \in \mathbb{R}^n : (\widetilde{Q^T x}) = \hat{A}^{-1} \tilde{y}_d \}$$

$$= \left\{ x \in \mathbb{R}^n : Q^T x = \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix} \text{ for some } z \in \mathbb{R}^{n-r} \right\}$$

$$\boxed{\text{i.e. } x = Q = \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix}}$$

$$= \left\{ Q \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix} : z \in \mathbb{R}^{n-r} \right\}$$

which global minimizer is best?
The global minimizers depend on z
Reasonable to use the shortest,
i.e. the shortest $Q \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix}$

Now:

$$\left\| Q \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ z \end{pmatrix} \right\|^2$$

since Q is orthogonal

$$= \| \hat{A}^{-1} \tilde{y}_d \|^2 + \| z \|^2$$

smallest for $z = 0$

Hence use the global minimizer

$$\hat{x} = Q \begin{pmatrix} \hat{A}^{-1} \tilde{y}_d \\ 0 \end{pmatrix},$$

which is the x of least norm which minimizes $\| y_d - Ax \|^2$.

Then:

$$\hat{x} = Q \begin{pmatrix} \hat{A}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{y}_d \\ \bar{y}_d \end{pmatrix} = P^T y_d$$

$$= \underbrace{Q \begin{pmatrix} \hat{A}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^T}_{A^+} y_d$$

Called A pseudo-inverse

$\Rightarrow$ The general definition of A^+ is :

(7.1) For $A \in \mathbb{R}^{m \times n}$

$A^+ \in \mathbb{R}^{n \times m}$ is defined so $A^+ y_d$ is the x of least norm which minimizes :

$$\| y_d - Ax \|^2 \text{ on } \mathbb{R}^n$$

Aside

For $A \in \mathbb{R}^{m \times n}$

$A^r \in \mathbb{R}^{n \times m}$ is a right inverse of A if $A A^r = I$

$A^l \in \mathbb{R}^{n \times m}$ is a left inverse of A if $A^l A = I$

A has a normal inverse iff A has a left inverse and a right inverse (and both are equal)

A might have an A^l or an A^r even when A^{-1} does not exist !

Some properties of A^+

-

A	an $\hat{x}$ minimizing $\| y_d - Ax \|^2$
A^{-1} exists	$A^{-1} y_d$
A^{-1} does not exist	$A^+ y_d$

$\Downarrow$

if A has a (normal) inverse : $A^+ = A^{-1}$

- if A has a right-inverse : A^+ is one
- if A has a left-inverse : A^+ is one
- if $A = 0$: $A^+ = 0$

$\Downarrow$

A^+ = a fantastic , do-it-all inverse

Some applications of A^+

① **Choice of input x so $Ax = y_d$**

choose $x = A^+ y_d$ $\boxed{\Rightarrow Ax \text{ is as close as possible to } y_d}$

If A has a right-inverse, then:

$$Ax = \underbrace{AA^+}_{I} y_d = y_d \qquad \text{NICE}$$

$\boxed{\text{since } A^+ \text{ is a right-inverse if } A \text{ has one}}$

② **Recovery of x from the measurement $y = Ax$**

choose $\tilde{x} = A^+ y$ $\Rightarrow$ $\tilde{x}$ is a good estimate of the x generating y

If A has a left-inverse: then:

$$\tilde{x} = A^+ y = \underbrace{A^+ A}_{I} x = x$$

$\boxed{\text{since } A^+ \text{ is a left-inverse if } A \text{ has one}}$

i.e. A^+ recovers x without error

Computation of A^+

The detailed structure of the orthogonal decomposition of A affects formula + props. for A^+

A	A^+	Properties
$P\begin{bmatrix}\hat{A} & 0\\ 0 & 0\end{bmatrix}Q^T$	$Q\begin{bmatrix}\hat{A}^{-1} & 0\\ 0 & 0\end{bmatrix}P^T$	$A^+A \neq I$ i.e. $A^+ \neq$ a l-inv. $AA^+ \neq I$ i.e. $A^+ \neq$ a r-inv. $\nexists A^{-1}$
$P[\hat{A}\ 0]Q^T$	$Q\begin{bmatrix}\hat{A}^{-1}\\ 0\end{bmatrix}P^T$	$A^+A \neq I$ $AA^+ = I$ i.e. $A^+ =$ a r-inv. $\nexists A^{-1}$
$P\begin{bmatrix}\hat{A}\\ 0\end{bmatrix}Q^T$	$Q[\hat{A}^{-1}\ 0]P^T$	$A^+A = I$ i.e. $A^+ =$ a l-inv. $AA^+ \neq I$ $\nexists A^{-1}$
$P\hat{A}Q^T$	$Q\hat{A}^{-1}P^T$	$A^+ = A^{-1}$

Hence: calculation of A^+ easy once the orthogonal decomposition of A is done