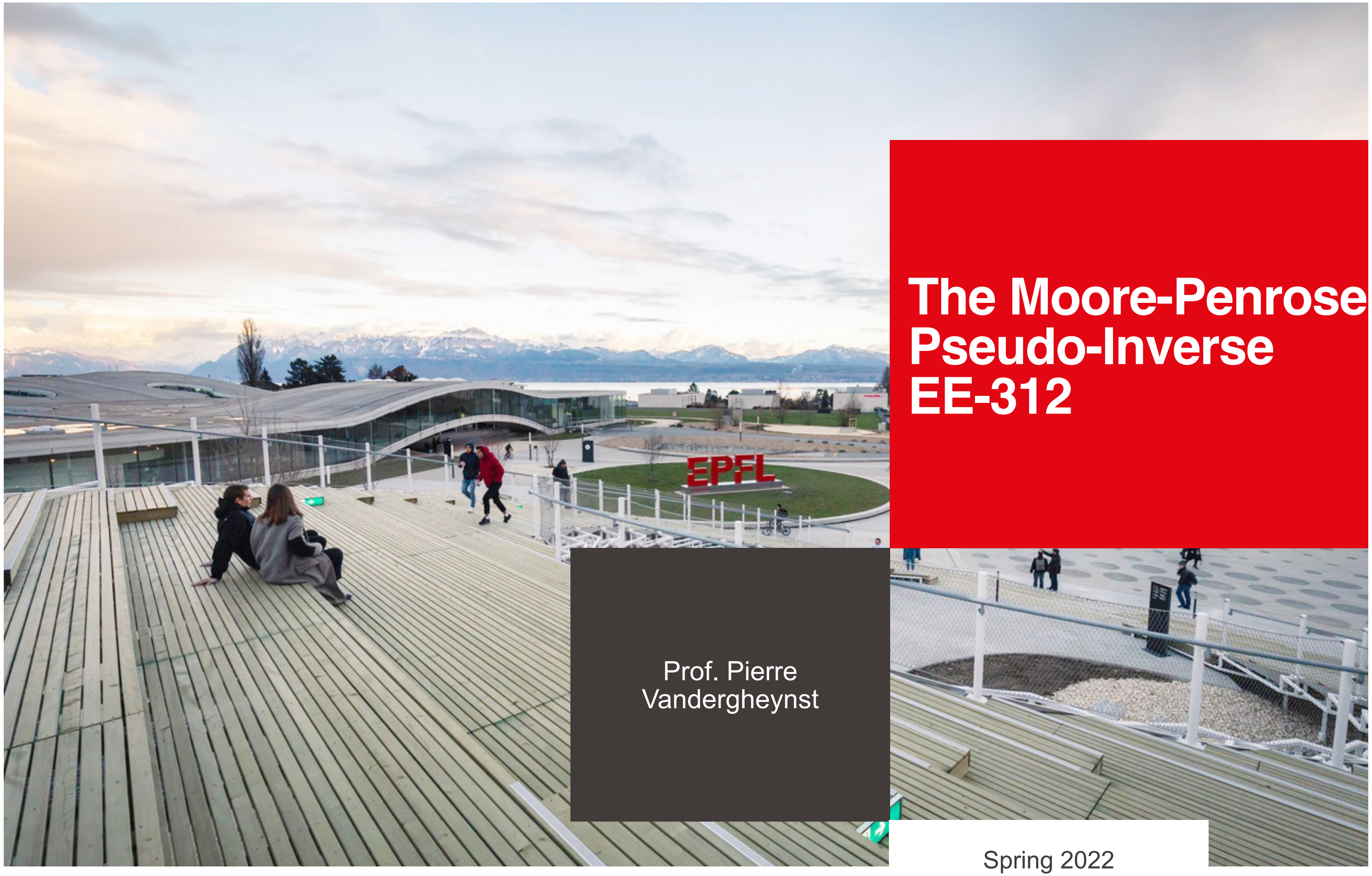
EPFL



The Moore-Penrose Pseudo-Inverse EE-312

Prof. Pierre Vandergheynst

Spring 2022

A note on invertibility

A linear transformation is **invertible** if and only if it is **bijective** (1-1 and onto)

ex: $T : \mathcal{N}(A)^{\perp} \to \mathcal{R}(A) \qquad Tv = Av \quad \forall v \in \mathcal{N}(A)^{\perp}$

Consider $A : \mathbb{R}^n \to \mathbb{R}^n$ and suppose it has n linearly independent columns $\mathcal{R}(A) = \mathbb{R}^n$

full-rank
A is onto

$$\forall y \in \mathbb{R}^n \, \exists \, x_1, ..., x_n \text{ s.t. } y = a_1 x_1 + \cdots + a_n x_n$$

unique
A is 1-1

$$y = Ax$$

x → depends on A and y

$$x = A^{-1} y$$

linear

> $A : \mathcal{V} \to \mathcal{W}$ is **invertible** if and only if it is **bijective.**
> If A is invertible then $\dim(\mathcal{V}) = \dim(\mathcal{W})$
> $A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible (non-singular) if and only if $\text{rank}(A) = n$

A note on invertibility

Let $A : \mathbb{R}^n \to \mathbb{R}^m$

Some interesting square matrices from any matrix $A^T A \in \mathbb{R}^{n \times n}$ and $AA^T \in \mathbb{R}^{m \times m}$

A is onto iff $\text{rank}(A) = m$, equivalent to AA^T is non-singular

intuition: range of A is full, null-space of A^T is trivial

so the induced linear transformation from $\mathbb{R}^m$ to itself is bijective

A is 1-1 iff $\text{rank}(A) = n$, equivalent to $A^T A$ is non-singular

intuition: range of A^T is full, null-space of A is trivial

so the induced linear transformation from $\mathbb{R}^n$ to itself is bijective

These matrices play a crucial role for constructing "special inverses"

A note on invertibility

Even if A is not invertible, it is left (resp. right) invertible iff it is 1-1 (resp. onto)

$A : \mathcal{V} \to \mathcal{W}$

A is **right invertible** if there exists a linear transformation

$$A_R^{-1} : \mathcal{W} \to \mathcal{V} \quad \text{such that } AA_R^{-1} = \mathbb{I}_{\mathcal{W}}$$

A is right invertible IFF it is onto

A is **left invertible** if there exists a linear transformation

$$A_L^{-1} : \mathcal{V} \to \mathcal{W} \quad \text{such that } A_L^{-1}A = \mathbb{I}_{\mathcal{V}}$$

A is left invertible IFF it is 1-1

A note on invertibility

A is invertible IFF it is both left and right invertible, in which case $A_L^{-1} = A_R^{-1} = A^{-1}$

Now our special square matrices become useful

A is onto $\Rightarrow AA^T$ is nonsingular $\Rightarrow A_R^{-1} = A^T(AA^T)^{-1}$ is a right inverse

A is 1-1 $\Rightarrow A^TA$ is nonsingular $\Rightarrow A_L^{-1} = (A^TA)^{-1}A^T$ is a left inverse

Rem: $A : \mathcal{V} \to \mathcal{V}$

if there exists a *unique* left inverse, then A is invertible

if there exists a *unique* right inverse, then A is invertible

Generalized Inverses

A little motivation (we'll get back to it later)

$$Ax = b \qquad A \in \mathbb{R}^{m\times m} \text{ and non-singular} \Rightarrow x = A^{-1}b$$

$A \in \mathbb{R}^{n\times m}$ and $b \in \mathcal{R}(A)$

Suppose there exists $G \in \mathbb{R}^{m\times n}$ s.t $AGy = y \quad \forall y \in \mathcal{R}(A)$

$x = Gb$ is a solution and $AGA = A$

Generalized Inverses

$$\forall A \in \mathbb{R}^{m \times n} \exists G \in \mathbb{R}^{n \times m} \text{ s.t } AGA = A$$

$\boldsymbol{G}$ = **Generalized Inverse** (always exists but not necessarily unique)

If the inverse of $A \in \mathbb{R}^{n \times n}$ exists, it is a generalised inverse (and there is only one)

$$A(A^{-1})A = A$$

$$\begin{aligned} G &= A^{-1}(AGA)A^{-1} \text{ (inverse exists)} \\ &= A^{-1}(A)A^{-1} \text{ (generalized inverse)} \\ &= A^{-1} \end{aligned}$$

Example: $A \in \mathbb{R}^{1\times 2}$ $A = [1, 2]$

compute G? $G \in \mathbb{R}^{2\times 1}$ $G = \begin{pmatrix} x \\ y \end{pmatrix}$ $x, y \in \mathbb{R}$

$$AGA = A \Rightarrow (1, 2)\begin{pmatrix} x \\ y \end{pmatrix}(1, 2) = (1, 2)$$

$$(x + 2y) \cdot (1, 2) = (1, 2)$$

$$x + 2y = 1 \qquad G = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$A A_R^{-1} = \mathbb{I}_{\mathbb{R}^m} \qquad \underline{A A_R^{-1}} A = A$$

Generalized Inverses

An explicit form for G:

$$A \in \mathbb{R}_r^{m\times n} \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$ with A_{11} is r-by-r and invertible

Then: $G = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n\times m}$ is a generalised inverse IFF $A_{22} = A_{21} A_{11}^{-1} A_{12}$

Any m-by-n matrix of rank r can be put in that form by performing rows and columns permutations.

This shows any matrix has a generalised inverse. Can you see why ?

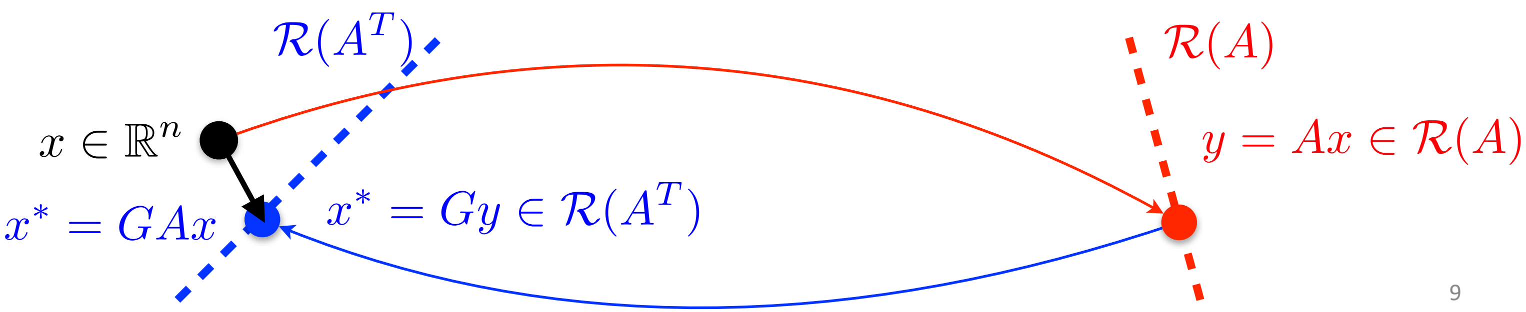
Generalised Inverses

A generalised inverse computes a projection !

$A \in \mathbb{R}^{m \times n}$ and $G \in \mathbb{R}^{n \times m}$ a g-inverse

AG is a projection onto $\mathcal{R}(A)$ column space

GA is a projection onto $\mathcal{R}(A^T)$ row space



$P = AG$ est une projection sur $R(A)$ $\quad A \in \mathbb{R}^{m \times n} \quad AG \in \mathbb{R}^{m \times m}$

$P^2 \overset{?}{=} P \qquad AG\,AG = (AGA)G$

$= AG = P\,!$

$R(P) \overset{?}{=} R(A)$

a. $y \in R(AG) \Rightarrow \exists x \in \mathbb{R}^m$ tq $y = AGx = A(Gx) \in R(A)$

$\Rightarrow R(AG) \subseteq R(A)$

b. $y \in R(A) \Rightarrow \exists x \in \mathbb{R}^n$ tq $y = Ax = AGAx$

$= AG(Ax) \in R(AG)$

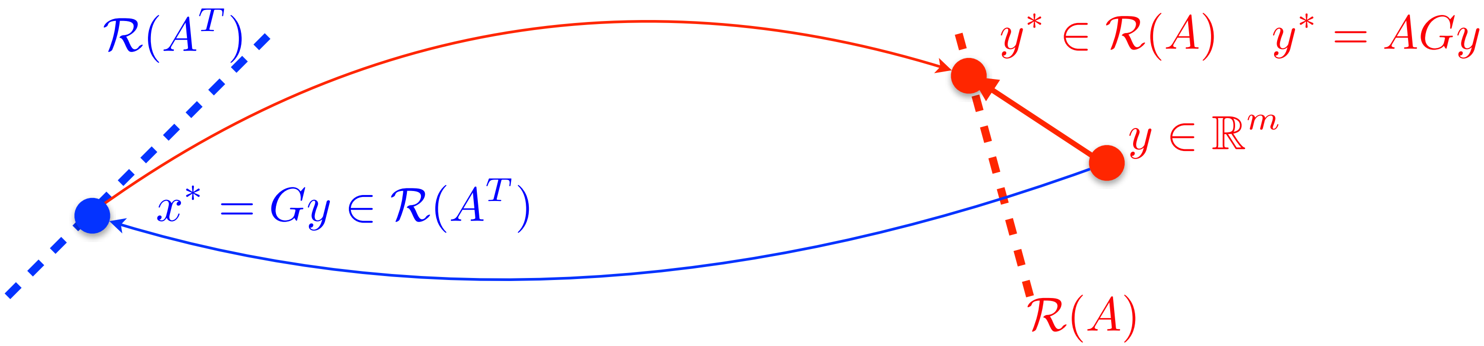
$\Rightarrow R(A) \subseteq R(AG)$

$\Rightarrow R(AG) = R(A)$

Generalised Inverses

Can this be useful to solve tougher problems ?

$A \in \mathbb{R}^{m\times n}$ and $G \in \mathbb{R}^{n\times m}$ a g-inverse $\quad Ax = y$ but $y \notin \mathcal{R}(A)$???



What would be the best possible y^* ?

The Moore-Penrose Pseudoinverse

Goal: show the existence and main properties of a generalised inverse for arbitrary matrices. Computational aspects and applications to systems of linear equations and least squares problem in next lectures

$A : \mathcal{X} \to \mathcal{Y}$ arbitrary linear transformation between finite dimensional vector spaces

$A^{\dagger}$

$T : \mathcal{N}(A)^{\perp} \to \mathcal{R}(A)$ $\quad Tx = Ax \quad \forall x \in \mathcal{N}(A)^{\perp}$ $\quad$ ***T* is bijective**

Define $A^{+} : \mathcal{Y} \to \mathcal{X}$ $\; A^{+}y = T^{-1}y_1$ where $y = y_1 + y_2$ with $y_1 \in \mathcal{R}(A)$ and $y_2 \in \mathcal{R}(A)^{\perp}$

A^{+} is the Moore-Penrose pseudoinverse of A

The Moore-Penrose Pseudoinverse

A^+ always exists and is unique! In particular for any rank r matrix $A \in \mathbb{R}_r^{m\times n}$

Some properties that characterise any $G = A^+$ **if and only if**

(P1) $AGA = A$

(P2) $GAG = G$

(P3) $(AG)^T = AG$

(P4) $(GA)^T = GA$

Rem: any nonsingular matrix satisfies (P1-4)

any left or right inverse satisfies at least 3 of these properties

The Moore-Penrose Pseudoinverse

Another characterisation

Let $A \in \mathbb{R}_r^{m \times n}$ then

$$A^+ = \lim_{\delta \to 0} (A^T A + \delta^2 \mathbb{I})^{-1} A^T$$
$$= \lim_{\delta \to 0} A^T (A A^T + \delta^2 \mathbb{I})^{-1}$$

Interestingly: (Proof: check properties P_1 – P_4

if A is onto (independent rows) then $A^+ = A^T (A A^T)^{-1}$ (right inverse)

if A is 1-to-1 (independent columns) then $A^+ = (A^T A)^{-1} A^T$ (left inverse)

Back to $A = (1, 2)$. We now $\exists$ ∞-many g-inverses $AGA = A$

ex: $G = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\qquad G = \begin{pmatrix} x \\ y \end{pmatrix}$

① $AGA = A \longrightarrow x + 2y = 1$

② $GAG = G \longrightarrow$ ok

③ $(AG)^T = AG \Longrightarrow x + 2y = x + 2y$

④ $(GA)^T = GA \Longrightarrow \begin{pmatrix} x \\ y \end{pmatrix}(1, 2) = \begin{pmatrix} x & 2x \\ y & 2y \end{pmatrix} = \begin{pmatrix} x & y \\ 2x & 2y \end{pmatrix}$

$x + 2y = 1$

$2x = y$

$x = \frac{1}{5}$

$y = \frac{2}{5}$

$A^+ = \begin{pmatrix} 1/5 \\ 2/5 \end{pmatrix}$

The Moore-Penrose Pseudoinverse

$I - AA^+ \rightarrow \mathcal{R}(A)^\perp$

It is a generalised inverse with more constraints!

It always exists (like g-inverse) and is unique (unlike most g-inverse)

(P1) $AGA = A$	G is a generalised inverse of A !
(P2) $GAG = G$	A is a generalised inverse of G !
(P3) $(AG)^T = AG$	AG is symmetric (and we know it is a projection ...)
(P4) $(GA)^T = GA$	GA is symmetric (and we know it is a projection ...)

$G = A^+$ $\quad (A^+)^+ = A$

AA^+ is the orthogonal projection onto $\mathcal{R}(A)$

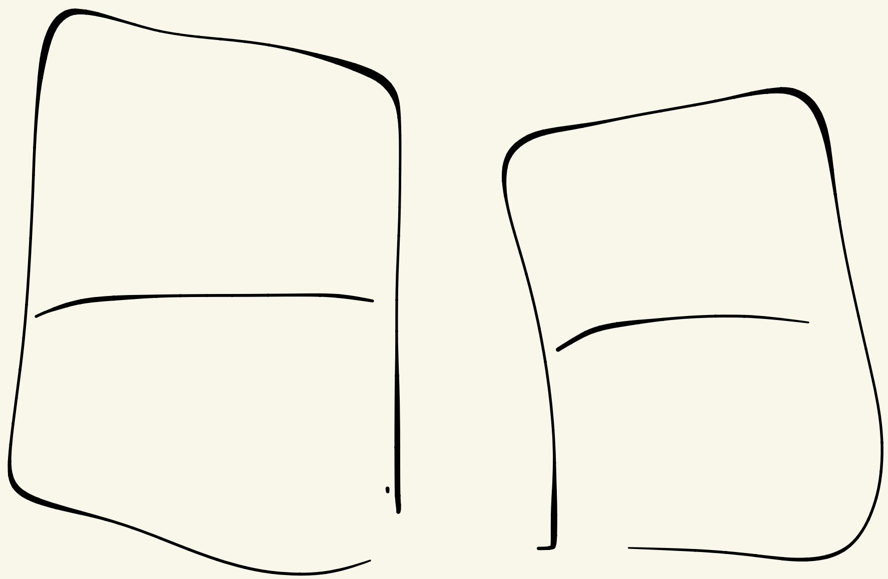
A^+A is the orthogonal projection onto $\mathcal{R}(A^T)$

$$P^{\perp}_{R(A)} = AA^{+}$$

$$P^{\perp}_{R(A)^{\perp}} = I - AA^{+}$$

$$P^{\perp}_{N(A)} = I - A^{+}A$$

$$P^{\perp}_{N(A)^{\perp}} = A^{+}A$$



The Moore-Penrose Pseudoinverse

$Ax = y$

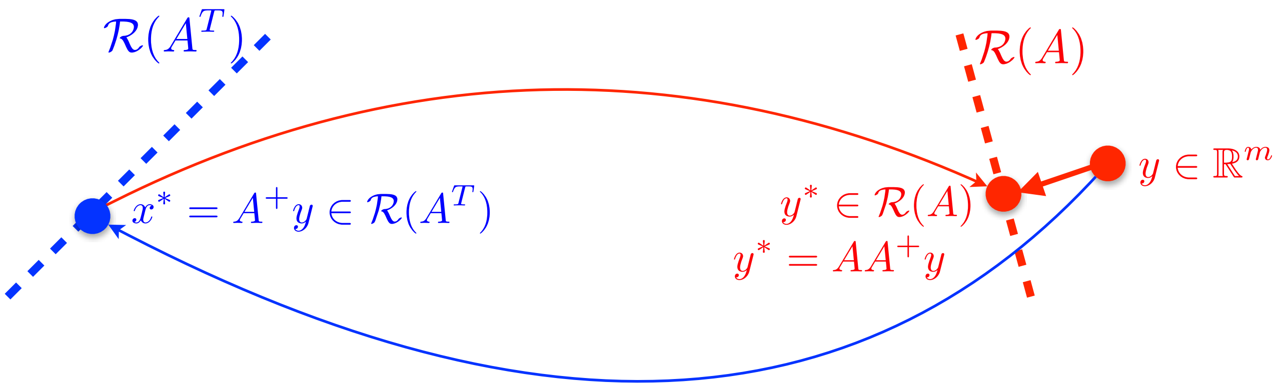
$Ax^* = y^*$

$\|y - y^*\|_2^2$

$A \in \mathbb{R}^{m \times n}$ is a tall matrix $m \geq n$, full column rank n

Ex: more equations than unknowns

$A^+ = (A^T A)^{-1} A^T$ we know this is invertible (full column rank)
and that the orthogonal projection on the range of A is: $P_{\mathcal{R}(A)} = A(A^T A)^{-1} A^T = AA^+$

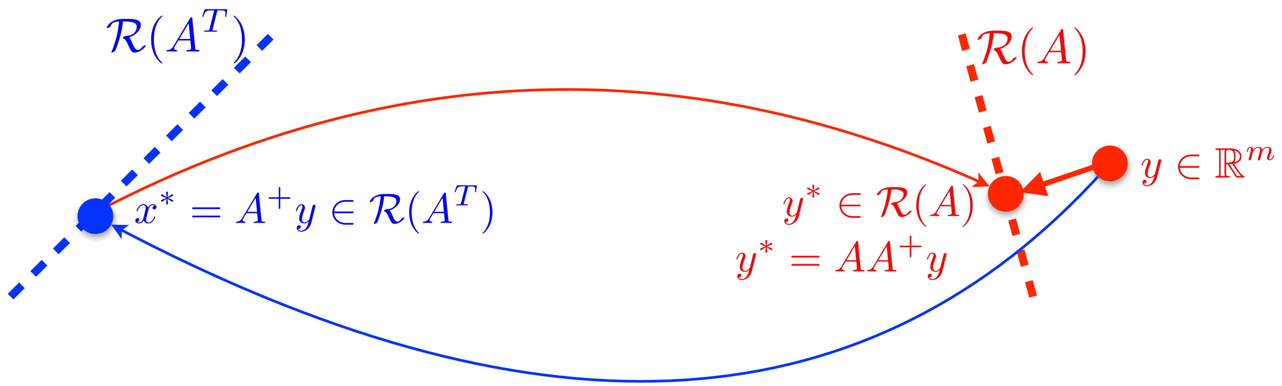


The Moore-Penrose Pseudoinverse

$A \in \mathbb{R}^{m \times n}$ is a tall matrix $m \geq n$, full column rank n

Ex: more equations than unknowns

$A^+ = (A^T A)^{-1} A^T$ we know this is invertible (full column rank) and that the orthogonal projection on the range of A is: $P_{\mathcal{R}(A)} = A(A^T A)^{-1} A^T = AA^+$



The Moore-Penrose Pseudoinverse

$A \in \mathbb{R}^{m \times n}$ is a tall matrix $m \geq n$, full column rank n

Ex: more equations than unknowns

$A^+ = (A^T A)^{-1} A^T$ we know this is invertible (full column rank) and that the orthogonal projection on the range of A is: $P_{\mathcal{R}(A)} = A(A^T A)^{-1} A^T = AA^+$