I. Linear Algebra

I.10. Pseudoinverse

Lecture based on

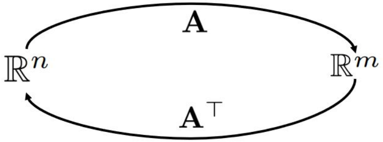
https://github.com/gwthomas/math4ml (Garrett Thomas, 2018)

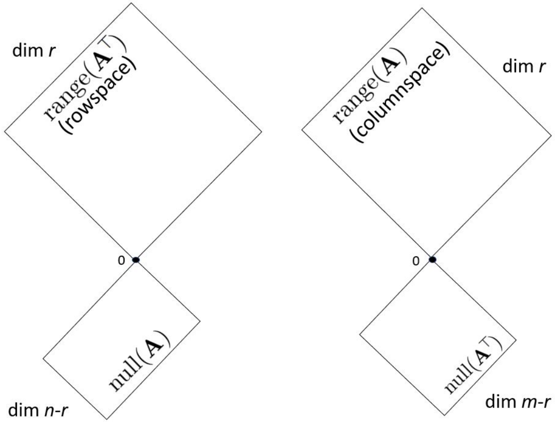
https://mml-book.github.io/ (Deisenroth et al., 2020, Mathematics for Machine Learning)

Prof. Dr. Christoph Lippert

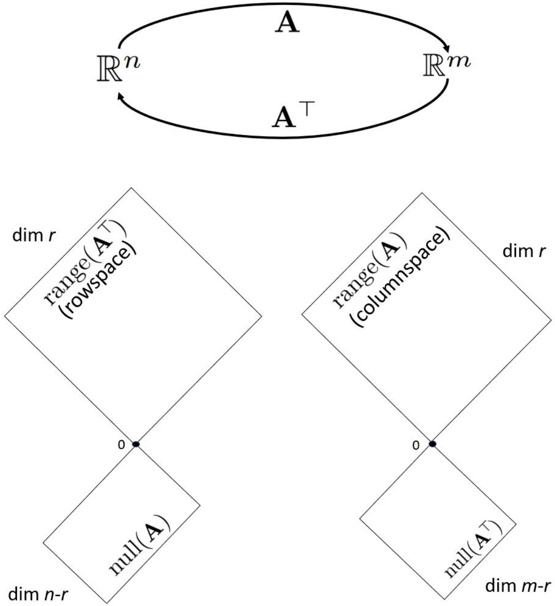
Digital Health & Machine Learning

$$\min_{\mathbf{x}} \frac{1}{2}(\mathbf{b}-\mathbf{A}\mathbf{x})^\top(\mathbf{b}-\mathbf{A}\mathbf{x})$$
$$\nabla(\mathbf{b}-\mathbf{A}\mathbf{x})^\top(\mathbf{b}-\mathbf{A}\mathbf{x}) = \mathbf{A}^\top(\mathbf{b}-\mathbf{A}\mathbf{x})$$





Moore-Penrose Pseudoinverse



The pseudo-inverse has the following properties:

i $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$

ii $\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$

iii It is unique.

$$\mathbb{R}^n \underset{\mathbf{A}^\top}{\overset{\mathbf{A}}{\rightleftarrows}} \mathbb{R}^m$$

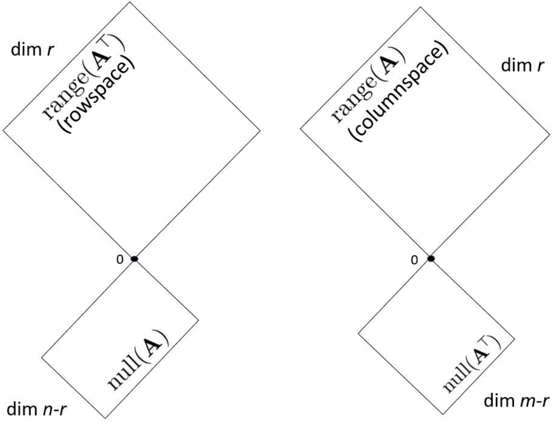
$$\mathbf{AV} = \mathbf{U\Sigma}$$

$$\mathbf{A}\mathbf{v}_1 = \sigma_1 \mathbf{u}_1 \quad \dots \quad \mathbf{A}\mathbf{v}_r = \sigma_r \mathbf{u}_r$$

$$\mathbf{A}\mathbf{v}_{r+1} = \mathbf{0} \quad \dots \quad \mathbf{A}\mathbf{v}_n = \mathbf{0}$$

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

$$\mathbf{A V_r} = \mathbf{U_r \Sigma_r}$$



Reduced Form of the SVD

$$\mathbf{A} = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_r & \dots & \mathbf{u}_n \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r & \dots & \mathbf{v}_n \end{bmatrix}^\top$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_r \end{bmatrix}^\top$$

$$\mathbf{A} = \mathbf{U_r \Sigma_r V_r}^\top$$