

0.0.1 The SVD solution for least squares

Here we mimic the development in the previous section and start with a matrix of full column rank: $\mathbf{A} \in \mathbb{C}_n^{m \times n}$. Therefore there is no null space component in the domain matrix. Because we have full column rank the null space for the domain is trivial and the general decomposition simplifies accordingly:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{\mathcal{R}} & \mathbf{X}_{\mathcal{N}} \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{X}_{\mathcal{R}} \end{bmatrix} \quad (0.1)$$

The goal is to convert the problem statement into a more tractable form exploiting unitary transformation. The minimization problem which defines the least squares solution is defined this way:

$$\|\mathbf{A}x - b\|_2^2 = \|\mathbf{Y}^* (\mathbf{A}x - b)\|_2^2 = \|\mathbf{Y}^* \mathbf{A}x - \mathbf{Y}^* b\|_2^2. \quad (0.2)$$

To simplify this and remove the \mathbf{A} matrix we need to rewrite the singular value decomposition:

$$\mathbf{A} = \mathbf{Y} \Sigma \mathbf{X}^* \quad \Rightarrow \quad \mathbf{Y}^* \mathbf{A} = \Sigma \mathbf{X}^*. \quad (0.3)$$

Using this relationship we can transform the norm as follows:

$$\|\mathbf{Y}^* \mathbf{A}x - \mathbf{Y}^* b\|_2^2 = \|\Sigma \mathbf{X}^* x - \mathbf{Y}^* b\|_2^2. \quad (0.4)$$

Expand then isolate range and null space contributions inside the norm

$$\begin{aligned} \|\Sigma \mathbf{X}^* x - \mathbf{Y}^* b\|_2^2 &= \left\| \left[\begin{array}{c|c} \mathbf{S} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \begin{bmatrix} \mathbf{X}_{\mathcal{R}}^* \\ \mathbf{X}_{\mathcal{N}}^* \end{bmatrix} x - \begin{bmatrix} \mathbf{Y}_{\mathcal{R}}^* \\ \mathbf{Y}_{\mathcal{N}}^* \end{bmatrix} b \right\|_2^2 \\ &= \left\| \begin{bmatrix} \mathbf{S} \mathbf{X}_{\mathcal{R}}^* x \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{Y}_{\mathcal{R}}^* b \\ \mathbf{Y}_{\mathcal{N}}^* b \end{bmatrix} \right\|_2^2. \end{aligned} \quad (0.5)$$

This result and the theorem of Pythagoras establishes the equivalence of the minimization problems

$$\min_{x \in \mathbb{R}^n} \|\mathbf{A}x - b\|_2^2 = \min_{x \in \mathbb{R}^n} \left(\|\mathbf{S} \mathbf{X}_{\mathcal{R}}^* x - \mathbf{Y}_{\mathcal{R}}^* b\|_2^2 + \|\mathbf{Y}_{\mathcal{N}}^* b\|_2^2 \right). \quad (0.6)$$

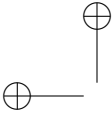
The classic definition of the least squares problem on the left has been reduced to an equivalent problem on the right which uses the fundamental subspace components on the left. This is a beautiful and powerful result which tells us two facts:

1. The minimizer x_{LS} solves

$$\mathbf{S} \mathbf{X}_{\mathcal{R}}^* x_{LS} = \mathbf{Y}_{\mathcal{R}}^* b. \quad (0.7)$$

2. The residual error r is again determined by the “silent” components of the codomain matrix

$$r^2 = \|\mathbf{Y}_{\mathcal{N}}^* b\|_2^2. \quad (0.8)$$



By minimizing the 2–norm of the error using the SVD expansion we have shown that the SVD solution for the least squares problem is given by the pseudoinverse:

$$x_{LS} = \mathbf{X}_{\mathcal{R}} \mathbf{S}^{-1} \mathbf{Y}_{\mathcal{R}}^* b = \mathbf{A}^+ b. \quad (0.9)$$

This is a remarkable result, for by simply using the singular value decomposition we have recovered the pseudoinverse solution! This is a natural consequence of the unitary transformation of the least squares problem. This is a main reason why the SVD and the pseudoinverse are considered to be intimately related subjects.

The norm of the error here is also given by the “silent” null space contribution

$$\|\mathbf{A} x_{LS} - b\|_2^2 = r^2 = \|\mathbf{Y}_{\mathcal{N}}^* b\|_2^2. \quad (0.10)$$

For a similar perspective see Golub, [?, p. 257].

Example: Once more we use the transpose of the example matrix $\mathbf{A} \in \mathbb{R}_2^{3 \times 2}$. The matrix and an singular value decomposition are given by this:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{Y}_{\mathcal{R}} & \mathbf{Y}_{\mathcal{N}} \end{bmatrix} \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{\mathcal{R}}^T \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 3 & 2 \\ 0 & 2 \end{bmatrix} &= \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} \\ \frac{5}{\sqrt{30}} & \frac{-1}{\sqrt{6}} & 0 \\ \frac{2}{\sqrt{30}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{15} & 0 \\ 0 & \sqrt{3} \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \end{aligned} \quad (0.11)$$

The data vector is

$$b = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}. \quad (0.12)$$

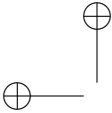
The least squares solution is then

$$\begin{aligned} x_{LS} &= \mathbf{X}_{\mathcal{R}} \mathbf{S}^{-1} \mathbf{Y}_{\mathcal{R}}^T b, \\ &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{15}} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{30}} & \frac{5}{\sqrt{30}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \\ &= \frac{1}{15} \begin{bmatrix} 1 \\ 6 \end{bmatrix}. \end{aligned} \quad (0.13)$$

This is the same answer as in equation (??).

The magnitude of the residual error vector is computed in this way:

$$r^2 = (\mathbf{Q}_{\mathcal{N}}^* b)^2 = \frac{16}{5}. \quad (0.14)$$



$$\begin{aligned} r^2 &= \left\| \mathbf{Y}_{\mathcal{N}}^T b \right\|_2^2, \\ &= \left\| \begin{bmatrix} \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\|_2^2, \\ &= \frac{16}{5}. \end{aligned} \tag{0.15}$$

This is in agreement with equation (??).