Deadline: Th. Jan. 24, 10:00 am Drop your printed or legible handwritten submissions into the boxes at Samelsonplatz.

Exercise 1 (SVD - 0 points)**.** Let A be a $m \times n$ matrix of rank r with SVD $A = U\Sigma V^\intercal$. Then $A^+ = V\Sigma^+U^\intercal$ is called it's *pseudo-inverse*, where

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0_{r\times n-r} \\ & & \sigma_r & \\ \hline & 0_{m-r\times r} & & 0_{m-r\times n-r} \end{array}\right] \qquad \Sigma^+ = \left[\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & 0_{r\times m-r} \\ & & \sigma_r^{-1} & \\ \hline & 0_{n-r\times r} & & 0_{n-r\times m-r} \end{array}\right]$$

In linear regression, we want to minimize $\|Y - X\beta\|_2^2$. We know that the optimal parameter β satisfies the normal equation $X^\intercal X\beta = X^\intercal Y$. Use the SVD to show that

- If $X^\intercal X$ is invertible, then $(X^\intercal X)^{-1}X^\intercal = X^+$
- $\beta = X^+Y$ solves the Normal equation, even if $X^\intercal X$ is not invertible

Exercise 2 (Q&A - 0 points)**.** Write up to 3 **short** questions you want to see discussed next week. We will select and prepare for some of the most often asked ones.