1. Problem

we consider

$$y = Ax \tag{6}$$

where $A \in R^{m\times n}$ is fat (m < n),$i.e.$,

- there are more variables than equations
- x is underspecified, $i.e.$, many choices of x lead to the same y

we'll assume that A is full rank(m), so for each $y \in R^m$, there is a solutions et of all solutions has form

$$\{x|Ax = y\} = \{x_p + z|Az = 0\} \tag{1}$$

where x_p is any ('particular') solution, $i.e.$, $Ax_p = y$

How to prove that $x^* = A^T(AA^T)^{-1}y$ is the solution of $y = Ax$ that minimizes $||x||_2^2$?

2. Proof

$x^* = A^T(AA^T)^{-1}y$ is a solution because

$$Ax^* = AA^T(AA^T)^{-1}y = Iy = y \tag{2}$$

Then we consider an optimization problem

$$\begin{aligned} &\max_x \ ||x||_2^2 \\ &s.t.\ Ax = b \end{aligned} \tag{3}$$

We have $A(x - x^*) = 0$ and $\forall x \in \{x|Ax = y\}$, so that

$$\begin{aligned} &(x - x^*)^T x^* = (x - x^*)^T A^T (AA^T)^{-1} y \\ &= (A(x - x^*))^T (AA^T)^{-1} y \\ &= 0 \end{aligned} \tag{4}$$

So $(x - x^*) \perp x^*$.

According to Triangular Inequality of 2 vertical vectors, we have

$$||x|| = \|x^* + x - x^*\|_2^2 = \|x - x^*\|_2^2 + \|x^*\|_2^2 \geq \|x^*\|_2^2 \tag{5}$$

So $x^* = A^T(AA^T)^{-1}y$ is the solution of $y = Ax$ that minimizes $||x||_2^2$