1. Problem

we consider

$$y = Ax \tag{1}$$

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

$$\{x|Ax = y\} = \{x_p + z|Az = 0\} \tag{2}$$

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 $||x||_2^2$?