

V.1 Linear Least-Squares

Let's focus on the linear case:

$$\min_{\vec{x}} \phi(\vec{x})$$

where $\phi(\vec{x}) = \frac{1}{2} \|A\vec{x} - \vec{b}\|_2^2$

and $A \in \mathbb{R}^{m \times n}$, $\vec{b} \in \mathbb{R}^m$, $\vec{x} \in \mathbb{R}^n$.

Moreover, we assume that A has full column rank n , i.e. the columns of A are linearly independent.

Have as usual: $m \dots$ number of measurements
 $n \dots$ number of parameters \downarrow

Rem.: The assumption on the rank implies that the parameters are not ambiguous

V.1.1 Normal equations

Let's rewrite $\phi(\vec{x})$ as:

$$\begin{aligned} \phi(\vec{x}) &= \frac{1}{2} \|A\vec{x} - \vec{b}\|_2^2 \\ &= \frac{1}{2} (A\vec{x} - \vec{b})^T (A\vec{x} - \vec{b}) = \frac{1}{2} (\vec{x}^T A^T - \vec{b}^T) (A\vec{x} - \vec{b}) \\ &= \frac{1}{2} \left(\vec{x}^T A^T A \vec{x} - \vec{x}^T A^T \vec{b} - \vec{b}^T A \vec{x} + \vec{b}^T \vec{b} \right) \\ &= \frac{1}{2} \left(\vec{x}^T A^T A \vec{x} - 2 \underbrace{\vec{x}^T A^T \vec{b}}_{\substack{= (\vec{b}^T A \vec{x})^T = \vec{x}^T A^T \vec{b} \\ \text{scalar!}}} + \vec{b}^T \vec{b} \right) \end{aligned}$$