Frage: Warum minimiert die Lösung des Normalgleichungssystems $A^T A \lambda = A^T y$ die Funktion $E(\lambda_1, \lambda_2, \dots, \lambda_m) = \|y - A\lambda\|_2^2$?

Antwort:

Um $\|y - A\lambda\|_2^2$ zu minimieren, müssen wir $\|y - A\lambda\|_2^2$ nach λ ableiten und die Ableitung gleich Null setzen (Es handelt sich um eine positive nach oben nicht beschränkte Funktion. Also ist ihr Extremum ein Minimum.)

$$\|y - A\lambda\|_2^2 = (y - A\lambda)^T \cdot (y - A\lambda)$$

$$\frac{\partial}{\partial\lambda}\|y - A\lambda\|_2^2 \overset{\text{Produktregel}}{=} \underbrace{\frac{\partial}{\partial\lambda}\left[(y - A\lambda)^T\right]}_{= -A^T} \cdot (y - A\lambda) +$$

$$+ (y - A\lambda)^T \cdot \underbrace{\frac{\partial}{\partial\lambda}(y - A\lambda)}_{= -A} = -A^T(y - A\lambda) -$$

$$-(y - A\lambda)^T A = -A^T y + A^T A\lambda - \underbrace{y^T A}_{= A^T y} + \underbrace{\lambda^T A^T A}_{= A^T A\lambda}$$

weil $A^T A$ symmetrisch ist

$$= 2A^T A\lambda - 2A^T y \overset{!}{=} 0$$

$$\Leftrightarrow A^T A\lambda = A^T y$$