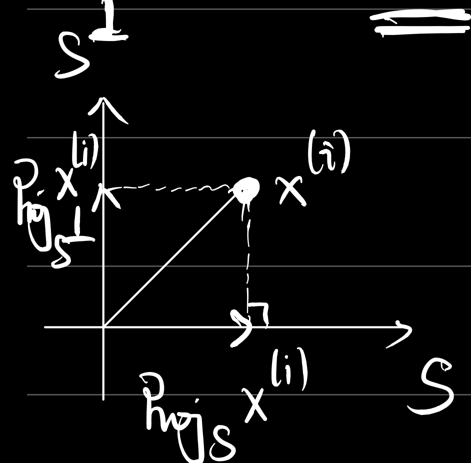
Problem Formulation 2: Given sample points $x^{(1)}, \dots, x^{(n)} \in \mathbb{R}^p$, find the d-dimensional subspace S with the least mean squared error.

Setting: Write $\{u^{(1)}, \dots, u^{(d)}\}$ for an orthonormal basis of S, the mean squared error

$$\overset{\text{derivation}}{=\!=} \sum_{i=1}^{n} \|x^{(i)}\|^2 - \sum_{j=1}^{d} \underbrace{u^{(j)T} X X^T u^{(j)}}_{\text{comes from } \|\mathrm{Proj}_S\, x^{(i)}\|^2}$$



It suffices to maximize $\sum_{j=1}^{d} u^{(j)T} X X^T u^{(j)}$

(a Rayleigh quotient)

derivation $\Downarrow$

Conclusion: $\{u^{(1)}, \dots, u^{(d)}\}$ are eigenvectors associated to the d largest eigenvalues of XX^T.

Problem Formulation 3: Given the sample matrix X, find a matrix with rank d that best approximate X w.r.t. the Frobenius norm, i.e.

$$\min_A \|X - A\|_F^2 \quad \text{s.t. } \operatorname{rank} A = d.$$

derivation $\Downarrow$ SVDs + von Neumann's inequality

Conclusion: If $X = U \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} V^T$ where $r = \operatorname{rank} X$

then

$$A = U \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_d & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} V^T$$

Related topics:
- def of Frobenius norm.
- properties of Frobenius norm.
- von Neumann's inequality.