Abteilung für
Angewandte Mathematik

Prof. Dr. Sören Bartels, Tatjana Schreiber June 4th, 2025

Algorithmic Aspects of Data Analytics and Machine Learning

SS 2025 — Sheet 7

https://aam.uni-freiburg.de/agba/lehre/ss25/algml/index.html

Due: June 20, 2025, 2 p.m.

Task 1 (5 Points)

Let $D = \{(3, 2, 2),\ (2, 3, -2)\} \subset \mathbb{R}^3$ and let

$$X = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$$

be the corresponding data matrix. Compute an orthonormal set of vectors $v_1, v_2, v_3 \in \mathbb{R}^3$ as follows:

1.) Find v_1 that maximizes $\|Xv\|$ under the constraint $\|v\| = 1$.

2.) Compute v_2 maximizing $\|Xv\|$ subject to $\|v\| = 1$ and $v \perp v_1$.

3.) Finally, determine v_3 as a unit vector maximizing $\|Xv\|$ subject to $\|v\| = 1$ and $v \perp v_1, v_2$.

The resulting vectors form an orthonormal basis aligned with the principal directions of the data.
Sketch the resulting subspaces spanned by $\{v_1\}$ and by $\{v_1, v_2\}$.
Hint: To maximize $\|Xv\|$ under the constraint $\|v\| = 1$, consider maximizing $\|Xv\|^2 = v^\top X^\top Xv$ instead. Introduce a Lagrange multiplier λ and solve

$$\nabla_v \left(v^\top X^\top Xv - \lambda(v^\top v - 1) \right) = 0.$$

Note: The subspaces spanned by $\{v_1\}$ and $\{v_1, v_2\}$ give the best rank-1 and rank-2 approximations to the data in terms of squared error. They are also referred to as 1-dimensional and 2-dimensional best-fitting subspaces.

Task 2 (4 Points)

Show that n data points in $\mathbb{R}^d$ do not, in general, uniquely determine a k-dimensional best-fitting subspace, even if the rank of the data matrix $X \in \mathbb{R}^{n\times d}$ is strictly greater than k.
We only consider the case $k = 1$ here. You can choose $D = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$ and maximize $\|Xv\|$ under the constraint $\|v\| = 1$ as described in Task 1.

Task 3 (3 Points)

Let $\mathbf{X} \in \mathbb{R}^{n\times s}$ and $\mathbf{Y} \in \mathbb{R}^{s\times m}$ be two matrices such that $\mathbf{X}$ and $\mathbf{Y}$ both have full rank s, with $s \le \min\{n, m\}$. Show that the matrix $\mathbf{A} = \mathbf{X} \cdot \mathbf{Y} \in \mathbb{R}^{n\times m}$ also has rank s.

Task 4 (4 points)

(i) Consider a graph G with Laplacian matrix L whose characteristic polynomial is given by

$$\det(L - \lambda I) = \lambda^3(\lambda - 2)(\lambda - 3)^4(\lambda - 4)^2.$$

How many connected components does the graph G have?

(ii) Determine the Cheeger constant of the following graph