

BIOSTAT M280, Homework 4

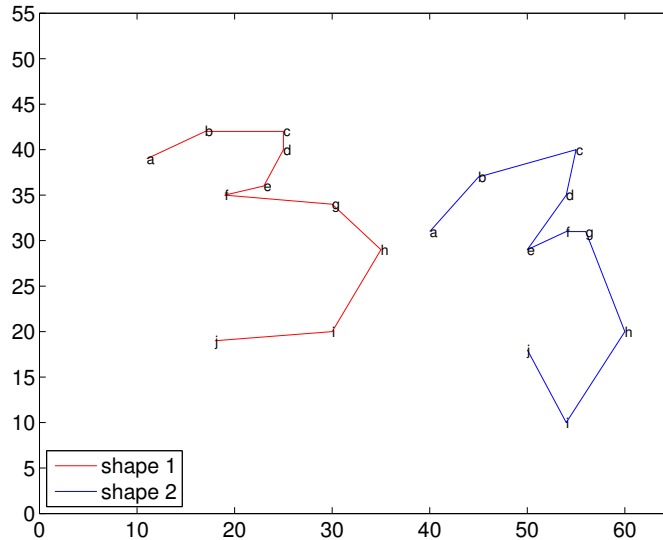
Due Tue, Feb 23 @ 11:59PM, 2016

1. (Ridge regression revisited) In ridge regression, we minimize

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_2^2,$$

where $\lambda \geq 0$ is a tuning parameter.

- (a) Express ridge solution $\hat{\boldsymbol{\beta}}(\lambda)$ in terms of the singular value decomposition (SVD) of \mathbf{X} .
 - (b) Show that (i) the ℓ_2 norms of ridge solution $\|\hat{\boldsymbol{\beta}}(\lambda)\|_2$ and corresponding fitted values $\|\hat{\mathbf{y}}(\lambda)\|_2 = \|\mathbf{X}\hat{\boldsymbol{\beta}}(\lambda)\|_2$ are non-increasing in λ and (ii) the ℓ_2 norm of the residual vector $\|\mathbf{y} - \hat{\mathbf{y}}(\lambda)\|_2$ is non-decreasing in λ .
 - (c) Re-compute and plot the ridge solution for the Longley data in HW3 at $\lambda = 5, 10, 15, 20, \dots, 100$ using the SVD approach.
 - (d) Comment on the computation efficiency of SVD approach compared to the approach you used in HW3.
2. (Matching images) Below figure displays two 3s written on a piece of paper and we want to properly align them.



Let matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times p}$ record n points on the two shapes. In this case $n = 10$ and $p = 2$. Mathematically we consider the problem

$$\text{minimize}_{\beta, \mathbf{O}, \boldsymbol{\mu}} \|\mathbf{X} - (\beta \mathbf{Y} \mathbf{O} + \mathbf{1}_n \boldsymbol{\mu}^T)\|_F^2,$$

where $\beta > 0$ is a scaling factor, $\mathbf{O} \in \mathbb{R}^{p \times p}$ is an orthogonal matrix, and $\boldsymbol{\mu} \in \mathbb{R}^p$ is a vector of shifts. Here $\|\mathbf{M}\|_F^2 = \sum_{i,j} m_{ij}^2$ is the squared Frobenius norm. Intuitively we want to rotate, stretch, and shift the shape \mathbf{Y} to match the shape \mathbf{X} as much as possible.

- (a) Let $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ be the column mean vectors of the matrices and $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ be the versions of these matrices with means removed. Show that the solution $(\hat{\beta}, \hat{\mathbf{O}}, \hat{\boldsymbol{\mu}})$ satisfies

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}} - \hat{\beta} \hat{\mathbf{O}}^T \bar{\mathbf{y}}.$$

Therefore we can center each matrix at its column centroid and then ignore the location completely.

- (b) Derive the solution to

$$\text{minimize}_{\beta, \mathbf{O}} \quad \|\tilde{\mathbf{X}} - \beta \tilde{\mathbf{Y}} \mathbf{O}\|_{\text{F}}^2$$

using the SVD of $\tilde{\mathbf{Y}}^T \tilde{\mathbf{X}}$.

- (c) Implement your method and solve the alignment problem in the figure. Display your solution together with the original two 3s.