Lecture 18:
Dimensionality Reduction

Professor Ilias Bilionis

Principal component analysis:
Basics

PREDICTIVE
SCIENCE LABORATORY

Principal component analysis as linear dimensionality reduction

$\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^D \longrightarrow \underline{z}_1, \underline{z}_2, \ldots, \underline{z}_n \in \mathbb{R}^d, \quad \underline{d \ll D}.$

① Projection

$\underline{x} \xrightarrow[\text{map}]{} \underline{z} = f(\underline{x}) = \underline{\underline{W}}^T \underline{x}, \quad \underline{\underline{W}} \in \mathbb{R}^{D\times d}$

principal components of $\underline{x}$

? – need to find

$(d\times D)(D\times 1) \to d\times 1$

$\underline{\underline{V}} \in \mathbb{R}^{D\times d}$

(affine map)

② Reconstruction

$\underline{\underline{V}}\,\underline{z} + \underline{x}_0 = g(\underline{z}) = \tilde{\underline{x}} \xleftarrow[\text{map}]{} \underline{z}$

? – need to find – ?

$\underline{x}_0 \in \mathbb{R}^D$

Parameters: $\underline{\underline{W}}, \underline{\underline{V}}, \underline{x}_0$?

projection matrix, reconstruction matrix, offset

Minimum-error formulation of principal component analysis

loss = reconstruction error $x_{1:n}$

= $\sum$ error in reconstruction of the projection of each observation

$$= \sum_{i=1}^{n} \| \underbrace{g}_{\text{reconstruction}}(\underbrace{f}_{\text{projection}}(\underline{x}_i)) - \underline{x}_i \|_2^2$$

- sum of square errors

$$= \sum_{i=1}^{n} \| \underline{\underline{V}} (\underline{\underline{W}}^T \underline{x}_i) + \underline{x}_0 - \underline{x}_i \|^2$$

min over $\underline{\underline{V}}, \underline{\underline{W}}, \underline{x}_0$