THE UNIVERSITY OF TENNESSEE KNOXVILLE

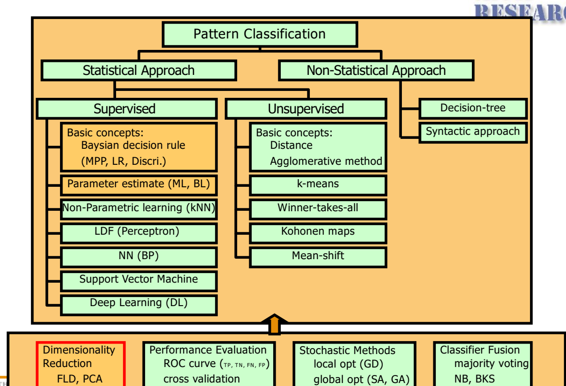
AICIP RESEARCH

ECE471-571 – Pattern Recognition

Lecture 7 – Dimensionality Reduction – Principal Component Analysis

Hairong Qi, Gonzalez Family Professor
Electrical Engineering and Computer Science
University of Tennessee, Knoxville
http://www.eecs.utk.edu/faculty/qi
Email: hqi@utk.edu

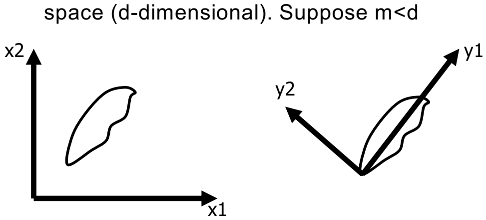
AICIP RESEARCH



AICIP RESEARCH

Principal Component Analysis or K-L Transform

- How to find a new feature space (m-dimensional) that is adequate to describe the original feature space (d-dimensional). Suppose m<d



THE UNIVERSITY OF TENNESSEE

3

K-L Transform (1)

AICIP RESEARCH

- Describe vector **x** in terms of a set of basis vectors $\mathbf{b}_i$.

$$\mathbf{x} = \sum_{i=1}^{d} y_i \mathbf{b}_i \qquad y_i = \mathbf{b}_i^T \mathbf{x}$$

- The basis vectors ($\mathbf{b}_i$) should be linearly independent and orthonormal, that is,

$$\mathbf{b}_i^T \mathbf{b}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

THE UNIVERSITY OF TENNESSEE 4

K-L Transform (2)

AICIP RESEARCH

- Suppose we wish to ignore all but *m* (*m*<*d*) components of **y** and still represent **x**, although with some error. We will thus calculate the first *m* elements of **y** and replace the others with constants

$$\mathbf{x} = \sum_{i=1}^{m} y_i \mathbf{b}_i + \sum_{i=m+1}^{d} y_i \mathbf{b}_i \approx \sum_{i=1}^{m} y_i \mathbf{b}_i + \sum_{i=m+1}^{d} \alpha_i \mathbf{b}_i$$

Error: $$\Delta \mathbf{x} = \sum_{i=m+1}^{d} (y_i - \alpha_i) \mathbf{b}_i$$

THE UNIVERSITY OF TENNESSEE 5

K-L Transform (3)

AICIP RESEARCH

- Use mean-square error to quantify the error

$$\begin{aligned}\varepsilon^2(m) &= E\left\{ \sum_{i=m+1}^{d} \sum_{j=m+1}^{d} (y_i - \alpha_i) \mathbf{b}_i^T (y_j - \alpha_j) \mathbf{b}_j \right\} \\ &= E\left\{ \sum_{i=m+1}^{d} \sum_{j=m+1}^{d} (y_i - \alpha_i)(y_j - \alpha_j) \mathbf{b}_i^T \mathbf{b}_j \right\} \\ &= \sum_{i=m+1}^{d} E\left\{ (y_i - \alpha_i)^2 \right\}\end{aligned}$$

THE UNIVERSITY OF TENNESSEE 6

K-L Transform (4)

AICIP RESEARCH

- Find the optimal α_i to minimize ε^2

$$\frac{\partial \varepsilon^2}{\partial \alpha_i} = -2\left(E\{y_i\} - \alpha_i\right) = 0$$

$$\alpha_i = E\{y_i\}$$

- Therefore, the error is now equal to

$$\varepsilon^2(m) = \sum_{i=m+1}^{d} E\left\{\left(y_i - E\{y_i\}\right)^2\right\}$$

$$= \sum_{i=m+1}^{d} E\left\{\left(\mathbf{b}_i^T \mathbf{x} - E\{\mathbf{b}_i^T \mathbf{x}\}\right)^2\right\} = \sum_{i=m+1}^{d} E\left\{\left(\mathbf{b}_i^T \mathbf{x} - E\{\mathbf{b}_i^T \mathbf{x}\}\right)\left(\mathbf{x}^T \mathbf{b}_i - E\{\mathbf{x}^T \mathbf{b}_i\}\right)\right\}$$

$$= \sum_{i=m+1}^{d} \mathbf{b}_i^T E\left\{\left(\mathbf{x} - E\{\mathbf{x}\}\right)\left(\mathbf{x} - E\{\mathbf{x}\}\right)^T\right\} \mathbf{b}_i = \sum_{i=m+1}^{d} \mathbf{b}_i^T \Sigma_{\mathbf{x}} \mathbf{b}_i = \sum_{i=m+1}^{d} \lambda_i$$

K-L Transform (5)

AICIP RESEARCH

- The optimal choice of basis vectors is the eigenvectors of $\Sigma_{\mathbf{x}}$
- The expansion of a random vector in terms of the eigenvectors of the covariance matrix is referred to as the Karhunen-Loeve expansion, or the "K-L expansion"
- Without loss of generality, we will sort the eigenvectors $\mathbf{b}_i$ in terms of their eigenvalues. That is $\lambda_1 >= \lambda_2 >= \ldots >= \lambda_d$. Then we refer to $\mathbf{b}_1$, corresponding to λ_1, as the "major eigenvector", or "principal component"

Summary

AICIP RESEARCH

- Raw data → covariance matrix → eigenvalue → eigenvector → principal component
- How to use error rate?