

Appendix A:

Principal Component Analysis

Principal Component Analysis (PCA) is almost equivalent to Singular Value Decomposition (SVD), or Karhunen-Loeve expansion. It will be presented first as an important computational method for feature extraction from input-data (Ritter, Martinetz & Schulten, 1992; Haken, 1996; AuxLit 10). To perform PCA, input-patterns \vec{x}^k are decomposed into a series, i.e. a linear combination of prototype-patterns \vec{w}^r ($r = 1, \dots, p'$):

$$\vec{x}^k = \vec{w}^0 + \sum_{r=1}^{p'} \vec{w}^r c_r(\vec{x}^k) + \vec{R}(\vec{x}^k) \quad (12.1)$$

DOI: 10.4018/978-1-61520-785-5.ch012

Copyright © 2011, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

$c_r = \vec{w}^r \vec{x}^k$ are the *principal components* of the input-patterns. Their number is p' . Variable c_r could also be treated as the activity-rate of corresponding dominant or cardinal neuron \vec{r} . Index r corresponds to location of a cardinal neuron \vec{r} . At the beginning, all principal components c_r are approximately equal and lower than 1. Then we can talk about potentially-cardinal neurons. Later environmental stimulus gives privilege to one pattern and its c_r increases towards 1 (wins). This means that one neuron becomes actually-cardinal, other neurons get subordinated. We are able to store all inputs \vec{x}^k completely (so, there is no need for data compression) if $p = p'$. In this “ideal” case index k and index r are equivalent. If, on the other hand, there is $p > p'$, data-compression causes that a higher number (p) of inputs \vec{x}^k is represented by a lower number (p') of cardinal neurons c_r or/and their set of synapses \vec{w}^r (or \vec{w}^r , respectively). c_r -representation is a relatively sparse code; \vec{w}^r is a relatively *distributed* code.

\vec{w}^0 specifies the center of weight of the input data distribution $P(\vec{x}^k)$ (\vec{w}^0 are very often represented by mean values – e.g., as in: Bankman in Pribram, 1993, p. 77). Other p' prototype-vectors \vec{w}^r form a so-called eigen-space basis. These basis-vectors, which are typically *mutually orthogonal, coincide with (relatively pictorial) memory representations*, i.e. so-called *eigen-images*. In fact, these p' vectors are eigen-vectors of the *autocorrelation-matrix* \mathbf{C} that have *the largest eigen-values* (Ritter et al., 1992):¹

$$\mathbf{C} = \sum_{k=1}^P (\vec{x}^k - \vec{w}^0) \otimes (\vec{x}^k - \vec{w}^0)^T P(\vec{x}^k). \quad (12.2)$$

T denotes the transposed vector (i.e., line vector); \otimes denotes the outer or tensor product: $(\vec{a} \otimes \vec{b})_{ij} = a_i b_j$. \mathbf{C} is an (auto) correlation matrix of the difference-vector $\vec{x}^k - \vec{w}^0$, therefore it is a *covariance matrix*. Typically, mutually orthogonal basis-vectors capture the directions of maximal variance (dispersion) of the data.

Equation (12.1) defines a hyper-plane which passes through the center of weight \vec{w}^0 and is spanned by principal axes along all \vec{w}^r . $\vec{R}(\vec{x}^k)$ is a residual vector which represents a non- vanishing distance from the approximating hyper-plane perpendicular to it. If $\vec{R}(\vec{x}^k)$ would be zero, our approximation with principal eigenvectors \vec{w}^r (prototypes) or, equivalently, with principal components c_r (cardinal neurons corresponding to prototypes) would be perfect.

2 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/appendix-principal-component-analysis/50511

Related Content

Evolutionary Search for Cellular Automata with Self-Organizing Properties toward Controlling Decentralized Pervasive Systems and Its Applications

Yusuke Iwase, Reiji Suzuki and Takaya Arita (2015). *International Journal of Systems Biology and Biomedical Technologies* (pp. 1-19).

www.irma-international.org/article/evolutionary-search-for-cellular-automata-with-self-organizing-properties-toward-controlling-decentralized-pervasive-systems-and-its-applications/148681

Classifier Ensembles Built on Subsets of Features

(2011). *Feature Selection and Ensemble Methods for Bioinformatics: Algorithmic Classification and Implementations* (pp. 260-295).

www.irma-international.org/chapter/classifier-ensembles-built-subsets-features/53908

Overview of Knowledge Discovery in Databases Process and Data Mining for Surveillance Technologies and EWS

Inci Batmaz and Guser Koksall (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 42-71).

www.irma-international.org/chapter/overview-knowledge-discovery-databases-process/76056

Polarization-based Robot Orientation and Navigation: Progress and Insights

Abd El Rahman Shabayek, Olivier Moreland and David Fofi (2015). *International Journal of Systems Biology and Biomedical Technologies* (pp. 73-89).

www.irma-international.org/article/polarization-based-robot-orientation-and-navigation/148684

Performance Evaluation: Final Check

(2011). *Feature Selection and Ensemble Methods for Bioinformatics: Algorithmic Classification and Implementations* (pp. 406-413).

www.irma-international.org/chapter/performance-evaluation-final-check/53915