Principal Components:

Introduction:

The principal component analysis (PCA) is used to:

a) transform correlated variables to uncorrelated ones,

b) search for linear combinations with large variability,

c) summarize high dimensional data,

d) reduce dimensionality of multivariate data.

Population Principal Components

Let $\mathbf{\Sigma}$ denote the covariance matrix of $p-$ dimensional random vector $\mathbf{x}$ with $\mu = E(\mathbf{x})$. Recall:

Spectral (Jordan) Decomposition:

Any symmetric $\mathbf{A}$ can be decomposed as: $\mathbf{\Gamma\Delta\Gamma}'$ where $\mathbf{\Delta}$ is a diagonal matrix of the eigenvalues of $\mathbf{A}$ and $\mathbf{\Gamma}$ is an orthogonal matrix whose columns are standardized eigenvectors.

Take $\mathbf{A} = \mathbf{\Sigma}$ (and using $\mathbf{\Gamma}$ orthogonal) we have:

$$\mathbf{\Gamma}'\mathbf{\Sigma}\mathbf{\Gamma} = \mathbf{\Delta},$$

where $\mathbf{\Delta}$ is a diagonal matrix of the (ordered) eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{\Gamma}$ is a $p \times p$ orthogonal matrix.

The principal component transformation is defined by:

$$\mathbf{y} = \mathbf{\Gamma}'(\mathbf{x} - \mu)$$

and the i'th component of $\mathbf{y} = (y_1, ..., y_p)$ is the i'th princical component (PC) of $\mathbf{x}$.

In particular, for each i:

$$y_i = \gamma_i'(\mathbf{x} - \mu),$$

where γ_i is the ith column of $\mathbf{\Gamma}$, the ith "Principal Component Loading"

e.g.

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \mathbf{\Sigma} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Largest Eigenvalue: 4

Eigenvector: $\gamma_1 = 2^{-1/2}(1, 1)'$

First PC: $y_1 = 2^{-1/2}(x_1 + x_2)$

Note: $E(y_1) = 0$ and $Var(y_1) = 4 = \lambda_1$

Smaller Eigenvalue: 2

Eigenvector: $\gamma_2 = 2^{-1/2}(1, -1)'$

Second PC: $y_2 = 2^{-1/2}(x_1 - x_2)$

Note: $E(y_2) = 0$ and $Var(y_2) = 2 = \lambda_2$.

Further: $Cov(y_1, y_2) = 0$

Always true?

Note that:

$$Var(\mathbf{y}) = \mathbf{\Gamma}'\mathbf{\Sigma}\mathbf{\Gamma} = \mathbf{\Delta}.$$

Hence:

i) $E(y_i) = 0$, $i = 1, ..., p$

ii) $Var(y_i) = \lambda_i$, $i = 1, ..., p$

iii) $Cov(y_i, y_j) = 0$, $i \neq j$

iv) $Var(y_1) \geq Var(y_2) \geq ... \geq Var(y_p)$

v) $\sum_{i=1}^{p} Var(y_i) = tr(\mathbf{\Sigma})$

vi) $\prod_{i=1}^{p} Var(y_i) = |\mathbf{\Sigma}|$

vii) For any $\mathbf{a}$ with $\mathbf{a}'\mathbf{a} = 1$ $Var(\mathbf{a}'\mathbf{x}) \leq \lambda_1$

viii) If $\mathbf{a}'\mathbf{x}$ is uncorrelated with the first k PC's of $\mathbf{x}$, then $Var(\mathbf{a}'\mathbf{x}) \leq \lambda_{k+1}$