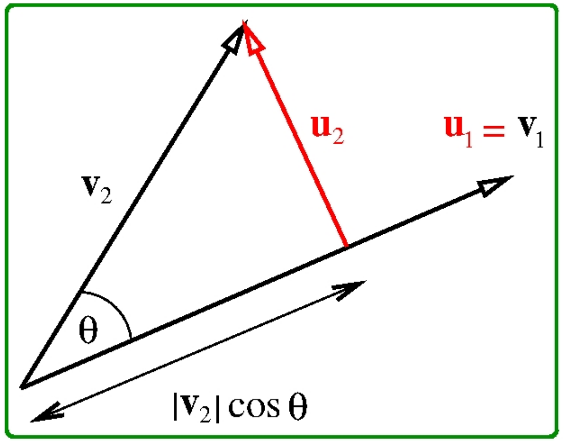
Gram-Schmidt Orthogonalisation

Orthonormalisation: compute all, then normalise!



- Dot product definition:
  $\mathbf{v_2} \cdot \mathbf{u_1} = |\mathbf{v_2}|\ |\mathbf{u_1}| \cos\theta$
  so $|\mathbf{v_2}| \cos\theta = \mathbf{v_2} \cdot \mathbf{u_1} / |\mathbf{u_1}|$
- This is the magnitude.
- Vector: magnitude $\times$ unit vector in that dirn
- unit vector in that dirn: $\mathbf{u_1}/|\mathbf{u_1}|$
- This vector: $\frac{\mathbf{v_2}\cdot\mathbf{u_1}}{|\mathbf{u_1}|} \frac{\mathbf{u_1}}{|\mathbf{u_1}|} = \frac{\langle \mathbf{v_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1}$
- Triangle law: this + $\mathbf{u_2}$ = $\mathbf{v_2}$
- Particular expression: $\boxed{\mathbf{u_2} = \mathbf{v_2} - \frac{\langle \mathbf{v_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1}}$
- General Expression: $\boxed{\mathbf{u_k} = \mathbf{v_k} - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v_k}, \mathbf{u_j} \rangle}{\langle \mathbf{u_j}, \mathbf{u_j} \rangle} \mathbf{u_j}}$

- Gram-Schmidt $>$ 2-D: graphical proof clumsy!
- Given $\mathbf{v}_1, \ldots \mathbf{v}_k$ linearly independent vectors (basis), to create an orthogonal set $\mathbf{u}_1, \ldots \mathbf{u}_k$
- Step 1: Start with $\mathbf{u}_1 = \mathbf{v}_1$
- Step 2: $(\perp)$ $\mathbf{u}_1$, $\mathbf{u}_2$ span same space as $\mathbf{v}_1, \mathbf{v}_2$
  Take $\mathbf{u_2} = a_1\mathbf{u}_1 + \mathbf{v}_2$ (lin combo, $\mathbf{u_1} = \mathbf{v_1}$)
  To find a_1, take a dot product with $\mathbf{u}_1$: (ortho)
  $a_1 = -\frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}$: $\boxed{\mathbf{u_2} = \mathbf{v_2} - \frac{\langle \mathbf{v_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle}\mathbf{u_1}}$
- Step 3: $(\perp)$ $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ span same space as $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$
  Take $\mathbf{u_3} = a_2\mathbf{u}_2 + a_1\mathbf{u}_1 + \mathbf{v}_3$ (lin combo, $\mathbf{v_1}, \mathbf{v}_2, \mathbf{v}_3$)
  a_1: take a dot product with $\mathbf{u}_1$: $a_1 = -\frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}$
  a_2: take a dot product with $\mathbf{u}_2$: $a_2 = -\frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}$
  $\boxed{\mathbf{u_3} = \mathbf{v_3} - \frac{\langle \mathbf{v_3}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle}\mathbf{u_1} - \frac{\langle \mathbf{v_3}, \mathbf{u_2} \rangle}{\langle \mathbf{u_2}, \mathbf{u_2} \rangle}\mathbf{u_2}}$
- General Expression: $\boxed{\mathbf{u_k} = \mathbf{v_k} - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v_k}, \mathbf{u_j} \rangle}{\langle \mathbf{u_j}, \mathbf{u_j} \rangle}\mathbf{u_j}}$

Home Page

Title Page

Contents

Page 39 of 55

Go Back

Full Screen

Close

Quit

- Eigenvectors of a symmetric matrix are orthonormal (Assumes no repeated eigenvalues) Actually follows from the diagonalisation result: $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ for symmetric matrices
  Explicit proof: consider $\mathbf{u}_i^T\mathbf{u}_j = (1/\lambda_j)\mathbf{u}_i^T\lambda_j\mathbf{u}_j = (1/\lambda_j)\mathbf{u}_i^T\mathbf{A}\mathbf{u}_j = (1/\lambda_j)(\mathbf{A}^T\mathbf{u}_i)^T\mathbf{u}_j = (1/\lambda_j)(\mathbf{A}\mathbf{u}_i)^T\mathbf{u}_j = (1/\lambda_j)(\lambda_i\mathbf{u}_i)^T\mathbf{u}_j = (\lambda_i/\lambda_j)\mathbf{u}_i^T\mathbf{u}_j \implies$ (as $\mathbf{A}^T = \mathbf{A}$) $(\lambda_i - \lambda_j)\mathbf{u}_i^T\mathbf{u}_j = 0 \implies \mathbf{u}_i \perp \mathbf{u}_j$, as $\lambda_i \neq \lambda_j$

Home Page
Title Page
Contents
Page 40 of 55
Go Back
Full Screen
Close
Quit

The K-L Transform

$$\mathbf{r}_i \triangleq \mathbf{U}^T \mathbf{p}_i \qquad \mathbf{R} \triangleq \mathbf{U}^T \mathbf{P}$$

$$\mathbf{r}_i = \begin{bmatrix} r_{i_1} \\ \vdots \\ r_{i_j} \\ \vdots \\ r_{i_k} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_j^T \\ \vdots \\ \mathbf{u}_k^T \end{bmatrix} \mathbf{p}_i$$

$\mathbf{r}_{ij} = \mathbf{p}_i \cdot \mathbf{u}_j \to \mathbf{p}_i$'s component along $\mathbf{u}_j$.

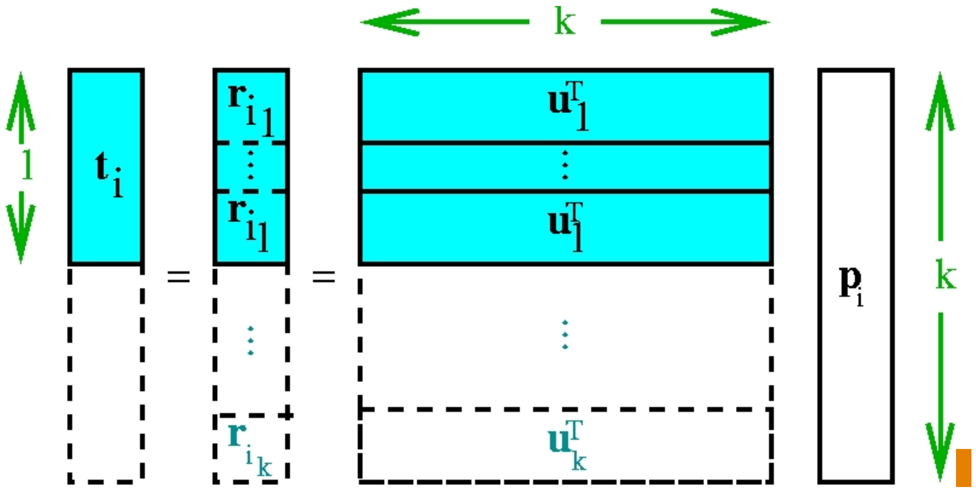
The Covariance Matrix

$$\hat{\mathbf{A}} \stackrel{\triangle}{=} \frac{1}{n}\,\mathbf{R}\,\mathbf{R}^T$$
$$= \frac{1}{n}\,\mathbf{U}^T\,\mathbf{P}\,(\mathbf{U}^T\,\mathbf{P})^T = \mathbf{U}^T\,\frac{1}{n}\,\mathbf{P}\,\mathbf{P}^T\,\mathbf{U}$$
$$= \mathbf{U}^T\,\mathbf{A}\,\mathbf{U} = \mathbf{U}^{-1}\,\mathbf{A}\,\mathbf{U} = \mathbf{\Lambda}\ \text{(Diagonalisation)}$$

In a Nutshell. . .

- Orthonormal matrix: Rotation connotation
- ‘Rotated’ patterns $\mathbf{r}_i$ line up with the EigenVectors
- The ‘rotated’ patterns are uncorrelated
- The average spreads of the ‘rotated’ patterns are the EigenValues themselves

Dimensionality Reduction



$$\boxed{\mathbf{t}_i \triangleq \mathbf{V}^T \mathbf{p}_i} \qquad \boxed{\mathbf{T} \triangleq \mathbf{V}^T \mathbf{P}}$$

$$\begin{aligned}\tilde{\mathbf{A}} &\triangleq \frac{1}{n}\mathbf{T}\,\mathbf{T}^T = \frac{1}{n}\mathbf{V}^T\,\mathbf{P}\,(\mathbf{V}^T\,\mathbf{P})^T \\ &= \mathbf{V}^T\,\mathbf{A}\,\mathbf{V} = \mathbf{V}^T\,\mathbf{U}\,\boldsymbol{\Lambda}\,\mathbf{U}^{-1}\,\mathbf{V} \text{ (Diagonalisation)} \\ &= (\mathbf{V}^T\,\mathbf{U})\,\boldsymbol{\Lambda}\,(\mathbf{V}^T\,\mathbf{U})^T = \boldsymbol{\Lambda}_l = diag(\lambda_1, \ldots\, \lambda_l)\end{aligned}$$

How many eigenvectors (l)?

e.g., min to make up 95% energy. $\min\, l : \frac{\sum_{i=1}^{l}\lambda_i}{\sum_{i=1}^{k}\lambda_i} \geq 0.95$