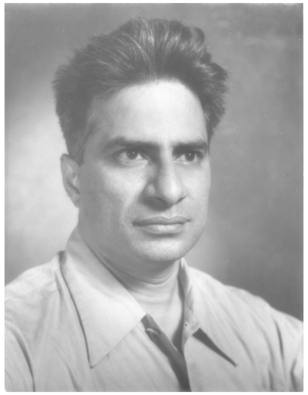
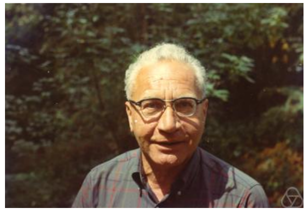
The ‘KLT’: Introduction

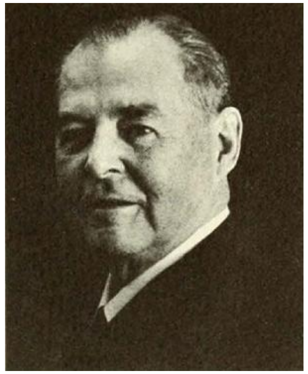


D. D. Kosambi
[1907-1966]
1943

K. Karhunen
[1915-1992]
1945



M. Loève
[1907-1979]
1948



H. Hotelling
[1895-1973]

https://upload.wikimedia.org/wikipedia/commons/0/0f/Kosambi-dd.jpg

https://upload.wikimedia.org/wikipedia/commons/d/d0/Michel_Lo%C3%A8ve.jpg

https://upload.wikimedia.org/wikipedia/en/4/49/Harold_Hotelling.jpg

People call it names!

- Karhunen-Loeve Transform
- Hotelling Transform
- Principal Component Analysis
- Eigenvalue-Eigenvector Transform

Home Page

Title Page

Contents

Page 28 of 55

Go Back

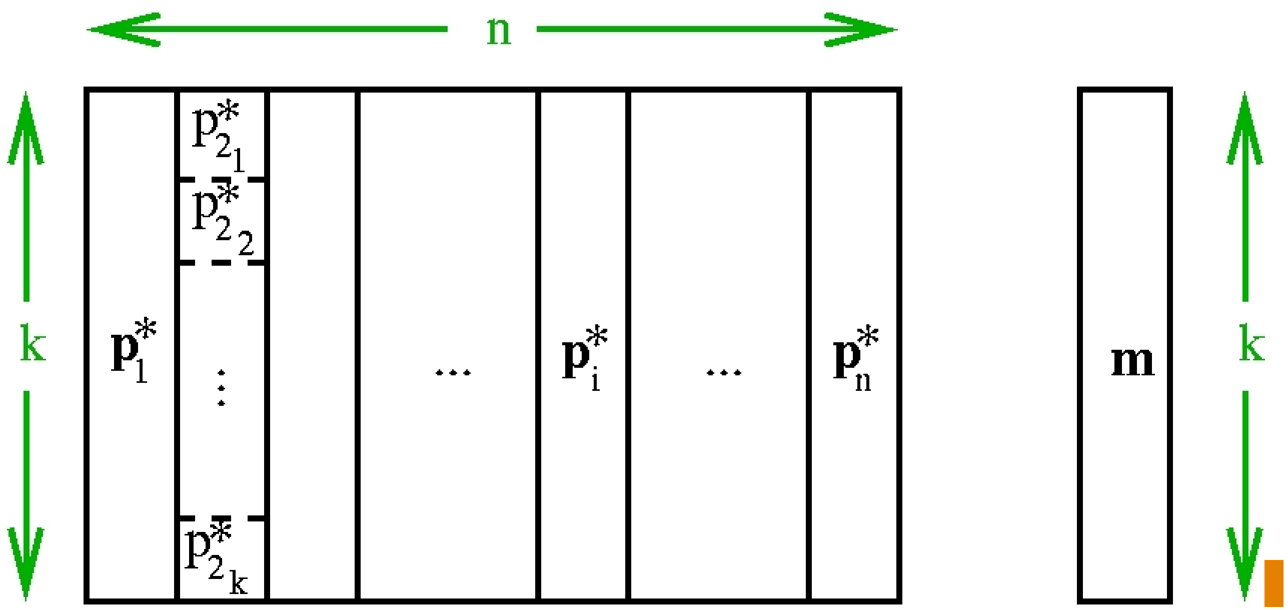
Full Screen

Close

Quit

Pattern Recognition Terms

- A 'pattern' is a $k \times 1$ column vector - a 1-D signal can be represented as a 'pattern'. A $k_1 \times k_2$ 2-D signal (an image) can be represented as a 'pattern' by taking all pixels in raster scan order (row major order) to form a $k \times 1$ 'pattern', $k = k_1 \cdot k_2$
- k-dimensional 'patterns' $\mathbf{p}_i^*$, $1 \leq i \leq n$
- Stack them up together (in any order) to form a $k \times n$ Pattern Matrix $\mathbf{P}^*$



- Normalise each pattern: $\mathbf{p}_i \triangleq \mathbf{p}_i^* - \mathbf{m}$
- $\mathbf{A} \triangleq \frac{1}{n}\,\mathbf{P}\,\mathbf{P}^T$: The Covariance Matrix
- Stack together EigenVectors $\mathbf{u}_i$ of $\mathbf{A}$ in decreasing order of the corresponding EigenValues to get the $k \times k$ matrix $\mathbf{U}$

Home Page

Title Page

Contents

Page 30 of 55

Go Back

Full Screen

Close

Quit

Linear Algebra Fundamentals

- Phys significance of Eigenvalues & Eigenvectors
- Similar Matrices
- Diagonalisation of a $k \times k$ matrix
- Gram-Schmidt Orthogonalisation
- Eigenvalues of a symmetric real matrix are real
- Eigenvecs of a symmetric matrix: orthonormality

Phys Sig of E'values, E'vectors

- For a $k \times k$ matrix $\mathbf{B}$, if $\mathbf{B}\mathbf{u_i} = \lambda_i \mathbf{u_i}$, λ_i are the eigenvalues, and $\mathbf{u_i}$, the corresponding eigenvectors
- Phys sig: matrix $\times$ vector $\equiv$ scaling it!
- Computing eigenvalues: $\mathbf{B}\mathbf{u} - \lambda\mathbf{u} = \mathbf{0} \Longrightarrow$ $(\mathbf{B} - \lambda\mathbf{I})\mathbf{u} = \mathbf{0} \Longrightarrow$ non-trival solution: $|\mathbf{B} - \lambda\mathbf{I}| = 0$
- E'vecs: not unique! scaled versions also e'vecs

Some Less Important Properties

- Rank($\mathbf{B}$) = # of non-zero eigenvals
- $\sum \lambda_i$ = Trace($\mathbf{B}$) (sum of main diag), $\prod \lambda_i = |\mathbf{B}|$
- A square matrix $\mathbf{A}$ and $\mathbf{A}^T$ have the same eigenvalues (but usually, different eigenvectors)
  $|\mathbf{A}^T - \lambda \mathbf{I}| = |\mathbf{A}^T - \lambda \mathbf{I}^T| = |(\mathbf{A} - \lambda \mathbf{I})^T| = |\mathbf{A} - \lambda \mathbf{I}|$
- The eigenvalues of a diagonal matrix are those!
  eigenvalues: $|\mathbf{B} - \lambda \mathbf{I}| = 0, \prod(b_{ii} - \lambda_i) = 0$
- $\mathbf{B}_{k \times k}$ is invertible iff 0 isn't an eigenvalue. eigenvalue 0 iff $|\mathbf{B} - 0\mathbf{I}| = 0$ iff $|\mathbf{B}| = 0$ i.e., non-invertible
- If $\mathbf{B}$ has an eigenvalue-eigenvector pair $(\lambda, \mathbf{u})$, then $\mathbf{B}^n$ $(n \in \mathcal{N})$ has the pair $(\lambda^n, \mathbf{u})$.
  $\mathbf{B}_{k \times k}\mathbf{u}_{k \times 1} = \lambda \mathbf{u}_{k \times 1}$, $\mathbf{B}\mathbf{B}\mathbf{u} = \lambda \mathbf{B}\mathbf{u}$, $\mathbf{B}^2\mathbf{u} = \lambda^2\mathbf{u}$, etc
- If $\mathbf{B}$ has an eigenvalue-eigenvector pair $(\lambda, \mathbf{u})$, then $\mathbf{B}^{-1}$ has the pair $(\lambda^{-1}, \mathbf{u})$.
  $\mathbf{B}_{k \times k}\mathbf{u}_{k \times 1} = \lambda \mathbf{u}_{k \times 1}$, $\mathbf{B}^{-1}\mathbf{B}\mathbf{u} = \lambda \mathbf{B}^{-1}\mathbf{u}$, $(1/\lambda)\mathbf{u} = \mathbf{B}^{-1}\mathbf{u}$

- Eigenvectors of a matrix with distinct eigenvalues are linearly independent: Can form a basis
  Proof by Contradiction: Suppose not. 'Thin out' this to l indep eigenvectors $\mathbf{u}_1, \ldots \mathbf{u}_l \equiv \lambda_1, \ldots \lambda_l$
  Suppose $\mathbf{u}$ was 'thinned out' $\mathbf{u} = \sum_{j=1}^{l} c_j \mathbf{u_j}$ (1)
  1. Multiply (1) by $\mathbf{B}$: $\mathbf{Bu} = \sum c_j(\mathbf{Bu_j})$, $\lambda \mathbf{u} = \sum c_j \lambda_j \mathbf{u_j}$
  2. Multiply (1) by λ: $\lambda \mathbf{u} = \sum c_j \lambda \mathbf{u}_j$
  Subtract: $\mathbf{0} = \sum c_j(\lambda - \lambda_j)\mathbf{u}_j$. Hence, $\forall j$:
  $c_j = 0$ (no!) or $\mathbf{u}_j = \mathbf{0}$ (no, as eigenvector is a non-trival solution) or $\lambda = \lambda_j$ (no!): Contradiction!

- Eigenvalues of a symmetric real matrix are real
  $\mathbf{Au} = \lambda \mathbf{u}$ and $\mathbf{A}^*\mathbf{u}^* = \lambda^*\mathbf{u}^*$, $\mathbf{A}^* = \mathbf{A}$: real
  Pre-multiply by $\mathbf{u}^{*T}$ and $\mathbf{u}^T$, and subtract:
  $\mathbf{u}^{*T}\mathbf{Au} - \mathbf{u}^T\mathbf{Au}^* = \lambda \mathbf{u}^{*T}\mathbf{u} - \lambda^*\mathbf{u}^T\mathbf{u}^*$ LHS: Consider $(\mathbf{u}^{*T}\mathbf{Au})^T$, scalar's transpose. $= \mathbf{u}^T\mathbf{Au}^*$. LHS = 0
  RHS: $\mathbf{u}^{*T}\mathbf{u}$: sum-of-sq $\neq 0$ unless all 0, e'vec $\neq \mathbf{0}$
  e.g., $[a - jb\ c - jd][a + jb\ c + jd]^T = a^2 + b^2 + c^2 + d^2$
  Hence $\lambda = \lambda^*$, only possible if λ is real. QED

Home Page
Title Page
Contents
Page 32 of 55
Go Back
Full Screen
Close
Quit

Home Page

Title Page

Contents

Page 33 of 55

Go Back

Full Screen

Close

Quit

- Definiteness of a symmetric matrix depends on the sign of its eigenvalues. Quadratic Form: for $\mathbf{B}_{k\times k}$ the scalar $\mathbf{x}^T\mathbf{B}\mathbf{x}$ is a quadratic form $= \sum_{i=1}^{k}\sum_{j=1}^{k} b_{ij}x_ix_j$. $(\mathbf{x}^T\mathbf{B}\mathbf{x})^T = \mathbf{x}^T\mathbf{B}^T\mathbf{x}$: $\mathbf{B}$ is symmetric!

  Quadratic form: defined only for symmetric $\mathbf{B}_{k\times k}$

  Positive Definite: $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0,\ \forall\mathbf{x}$

  Positive Semi-Definite: $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0,\ \forall\mathbf{x}$

  $(*)$ PSD matrix has non-negative eigenvalues:

  Let λ be an eigenvalue of $\mathbf{A}_{k\times k}$ with eigenvector $\mathbf{u}$. $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$: $\mathbf{u}^T\mathbf{A}\mathbf{u} = \lambda\mathbf{u}^T\mathbf{u}$. Since $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0\ \forall\mathbf{x}$, $\mathbf{u}^T\mathbf{A}\mathbf{u} \geq 0$. $\mathbf{u}^T\mathbf{u} \geq 0 \implies \lambda \geq 0$

  $(*)$ Non-negative eigenvalues $\implies$ PSD:

  Symmetric matrix $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^T$. $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{U}\Lambda\mathbf{U}^T\mathbf{x} = \mathbf{y}^T\Lambda\mathbf{y} = \sum_{i=1}^{k}\lambda_i y_i^2$. If all $\lambda_i \geq 0$, $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is PSD

- The Quadratic form for Constrained Optimisation of a Symmetric Matrix: $\max / \min_{\mathbf{x}} \mathbf{x}^T \mathbf{A}\mathbf{x}, \; ||\mathbf{x}||^2 = 1$

  Take $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \lambda(\mathbf{x}^T\mathbf{x} - 1)$ (or $(1 - \mathbf{x}^T\mathbf{x})$)

  $\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = 0 :$ $2\mathbf{A}\mathbf{x} + \lambda 2\mathbf{x} = \mathbf{0}$: $(\mathbf{A} + \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$

  For this, take $\mu = -\lambda$ (λ suffices for $(1 - \mathbf{x}^T\mathbf{x})$)

  $(\mathbf{A} - \mu \mathbf{I})\mathbf{x} = \mathbf{0}$: Soln: eigenvec $\mathbf{u}$ of $\mathbf{A} \equiv$ eigenval μ

  To optimise: $\mathbf{x}^T\mathbf{A}\mathbf{x}$. At the opt: $\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{u}^T\mathbf{A}\mathbf{u} = \lambda \mathbf{u}^T\mathbf{u} = \lambda \implies$ $\max = \lambda_{\max}, \; \min = \lambda_{\min}$

Home Page

Title Page

Contents

◀◀ ▶▶

◀ ▶

Page 35 of 55

Go Back

Full Screen

Close

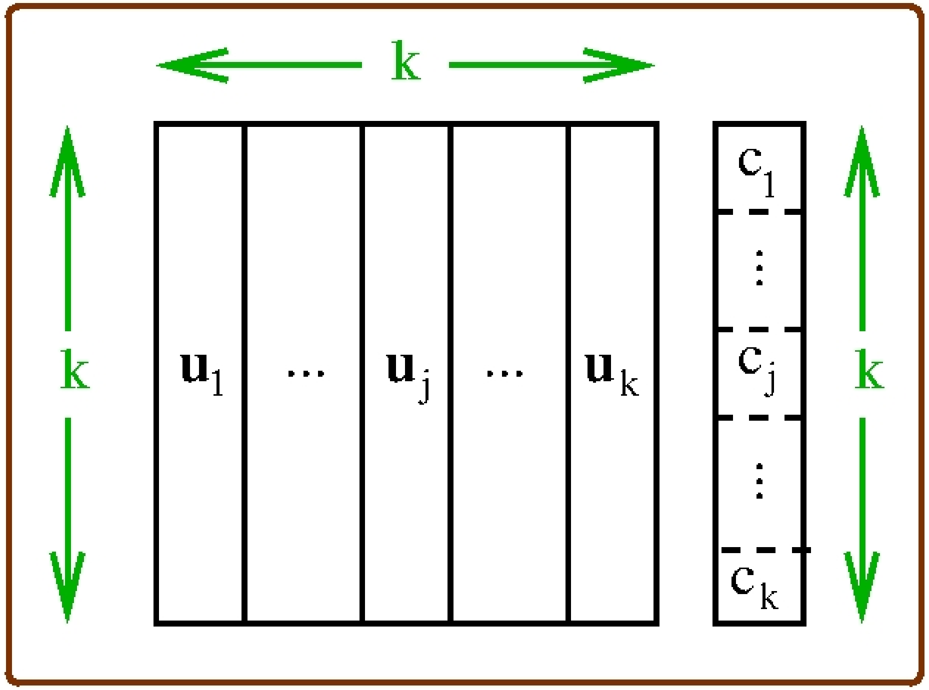
Quit

Similar Matrices

- For a $k \times k$ matrix $\mathbf{B}$ and any invertible $k \times k$ matrix $\mathbf{E}$, $\mathbf{EBE}^{-1}$ and $\mathbf{B}$ are Similar Matrices
- $\mathbf{Bu} = \lambda \mathbf{u}, \implies \mathbf{EBu} = \lambda \mathbf{Eu}, \implies \mathbf{EB}\ \mathbf{E}^{-1}\mathbf{E}\ \mathbf{u} = \lambda \mathbf{Eu}$ $\implies \mathbf{EBE}^{-1}\ \mathbf{Eu} = \lambda\ \mathbf{Eu}, \implies \mathbf{EBE}^{-1}\ \mathbf{v} = \lambda\ \mathbf{v}$

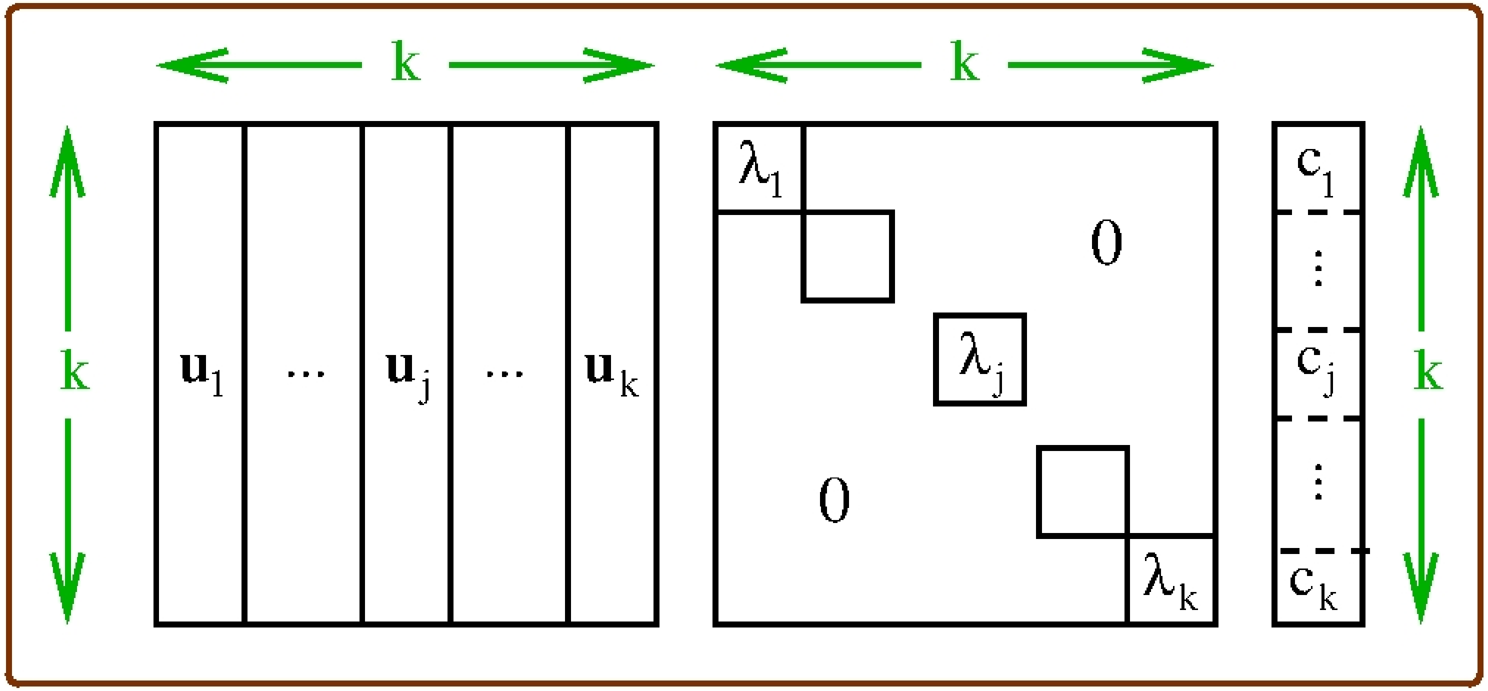
Diagonalisation of a $k \times k$ matrix $\mathbf{B}$

- $\mathbf{B}$: eigenvectors $\mathbf{u_i}$, stacked to get $\mathbf{U}$. $\mathbf{u_i}$'s not necessarily orthonormal, assume lin indep (non-repeated e'vals) $\implies$ basis of $k-$dim space
- What if accidentally end up with repeated eigen-values? SVD always works. Applications: use eigenvectors as an orthonormal basis. Extend with extra orthonormal vectors to span the space
- Any k-dimensional pattern $\mathbf{p_i}$ can be written as a linear combination of these basis vectors
- $\mathbf{p_i} = \sum_{j=1}^{k} c_j \mathbf{u_j}$



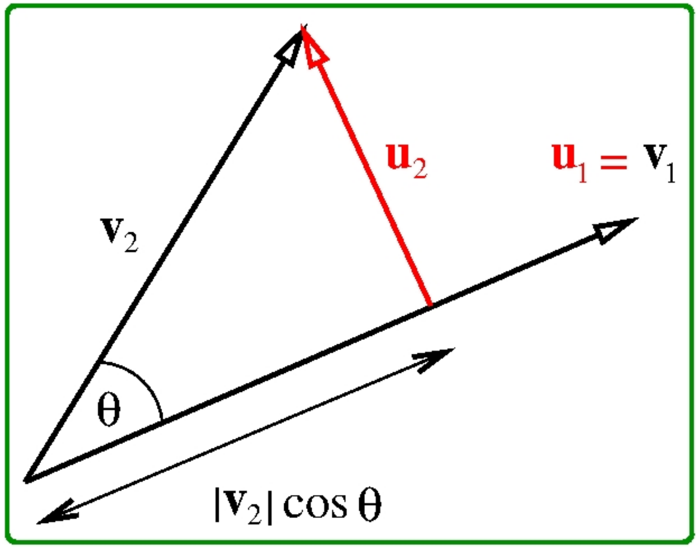
- $\mathbf{p_i} = \sum_{j=1}^{k} c_j \mathbf{u_j} = \mathbf{Uc}$
- $\mathbf{Bp_i} = \sum_{j=1}^{k} c_j \mathbf{Bu_j}$
  $= \sum_{j=1}^{k} c_j \lambda_j \mathbf{u_j} = \mathbf{U\Lambda c}$
- $\mathbf{BUc} = \mathbf{U\Lambda c} \implies$
  $\mathbf{BU} = \mathbf{U\Lambda} \implies$
- $\mathbf{B} = \mathbf{U\Lambda U}^{-1} \iff$
- $\mathbf{U}^{-1}\mathbf{BU} = \mathbf{\Lambda}$

If $\mathbf{U}$ is additionally orthonormal, $\mathbf{U}^{-1} = \mathbf{U}^T$



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 38 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 $\lambda_i \neq \lambda_j$