Math Tools: Linear Algebra

COMPSCI 773 S1 T

VISION GUIDED CONTROL

A/P Georgy Gimel'farb

Basic Notation

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$ - a rectangular m (rows) $\times$ n (columns) matrix

If $A_{ij} = 0$ for $i \neq j$ - a diagonal matrix $\mathrm{diag}\{A_{11}, A_{22}, \ldots, A_{kk}, \ldots\}$

I - the identity matrix: $I = \mathrm{diag}\{1, 1, \ldots, 1\}$; i.e. $I_{ii} = 1$; $I_{ij} = 0$ for $i \neq j$

$\mathbf{b} = [b_1, \ldots, b_m]^{\mathrm{T}}$ - a vector-column ($m \times 1$)

T - transposition of a matrix or a vector

$\mathbf{a} \bullet \mathbf{b} = a_1 b_1 + \ldots + a_m b_m$ - the dot product of two vectors

$\mathbf{a} \bullet \mathbf{b} = 0$ for the orthogonal vectors $\mathbf{a}$ and $\mathbf{b}$

The orthogonal vectors are **orthonormal** if simultaneously $\mathbf{a} \bullet \mathbf{a} = \mathbf{b} \bullet \mathbf{b} = 1$

Eigenvalues and Eigenvectors

- Eigenvalues and eigenvectors of a matrix A - solutions of the matrix-vector equation $A\mathbf{x} = \lambda\mathbf{x}$ or $(A - \lambda I)\mathbf{x} = 0$
  - Determinant $|A - \lambda I| = 0$ - the algebraic equation to find the eigenvalues
- If A is real-valued and symmetric $n \times n$ matrix, $A_{ij}=A_{ji}$, it has n mutually orthogonal eigenvectors $\mathbf{e}_1,\ldots,\mathbf{e}_n$
  - Typically, the eigenvectors are normalised to be orthonormal
  - Each eigenvector $\mathbf{e}_i$ has its own eigenvalue λ_i: $A\mathbf{e}_i = \lambda_i\mathbf{e}_i$

Example: 2 × 2 Symmetric Matrix

- $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ - the equation: $\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2 = 0$
- Eigenvalues:

$$\lambda_{1,2} = \tfrac{1}{2}(a+c) \pm \sqrt{\tfrac{1}{4}(a+c)^2 - ac + b^2} = \tfrac{1}{2}(a+c) \pm \sqrt{\tfrac{1}{4}(a-c)^2 + b^2}$$

- Identity matrix:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0 \Rightarrow \lambda_{1,2} = 1$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} e_{i,1} \\ e_{i,2} \end{bmatrix} = \begin{bmatrix} e_{i,1} \\ e_{i,2} \end{bmatrix}; \quad \begin{matrix} e_{i,1}^2 + e_{i,2}^2 = 1; \quad i = 1,2 \\ e_{1,1}e_{2,1} + e_{1,2}e_{2,2} = 0 \end{matrix} \quad \Rightarrow e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Example: 2 × 2 Symmetric Matrix

- For the identity matrix, the equations $A\mathbf{e}_i = \lambda_i\mathbf{e}_i$ do not constrain the eigenvectors (only identities $e_{i,j} = e_{i,j}$; $i,j = 1,2$)
  - Infinitely many pairs: $\mathbf{e}_1 = [\cos\theta, \sin\theta]^{\mathsf{T}}$; $\mathbf{e}_2 = [-\sin\theta, \cos\theta]^{\mathsf{T}}$
  - E.g. $\theta = 0$ for $\mathbf{e}_1 = [1,0]^{\mathsf{T}}$; $\mathbf{e}_2 = [0,1]^{\mathsf{T}}$

- $$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda_{1,2} = \pm 1$$

$$\lambda_1 = 1 \Rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} e_{1,1} \\ e_{1,2} \end{bmatrix} = \begin{bmatrix} e_{1,1} \\ e_{1,2} \end{bmatrix}; \quad \begin{matrix} e_{1,1}^2 + e_{1,2}^2 = 1 \\ e_{1,1} = e_{1,2} \end{matrix} \quad \Rightarrow e_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\lambda_2 = -1 \Rightarrow \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} e_{2,1} \\ e_{2,2} \end{bmatrix} = -\begin{bmatrix} e_{2,1} \\ e_{2,2} \end{bmatrix}; \quad \begin{matrix} e_{2,1}^2 + e_{2,2}^2 = 1 \\ e_{2,1} = -e_{2,2} \end{matrix} \quad \Rightarrow e_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Example: 3 × 3 Symmetric Matrix

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = 0 \Rightarrow -\lambda\left(\lambda^2 - 1\right) = 0 \Rightarrow \lambda_{1,2,3} = -1,0,1$$

$$\lambda_1 = 1 \Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} e_{1,1} \\ e_{1,2} \\ e_{1,3} \end{bmatrix} = \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ e_{1,3} \end{bmatrix} \Rightarrow e_{1,1} = 0; \quad e_{1,2} = e_{1,3} \Rightarrow e_1 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\lambda_2 = 0 \Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} e_{2,1} \\ e_{2,2} \\ e_{2,3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow e_{2,2} = e_{2,3} = 0; e_{2,1} = 1 \Rightarrow e_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_3 = -1 \Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} e_{3,1} \\ e_{3,2} \\ e_{3,3} \end{bmatrix} = -\begin{bmatrix} e_{3,1} \\ e_{3,2} \\ e_{3,3} \end{bmatrix} \Rightarrow e_{3,1} = 0; \quad e_{3,2} = -e_{3,3} \Rightarrow e_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Example: 3 × 3 Symmetric Matrix

$$A = \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix} \Rightarrow \begin{vmatrix} a-\lambda & b & 0 \\ b & a-\lambda & b \\ 0 & b & a-\lambda \end{vmatrix} = 0 \Rightarrow (a-\lambda)\left((a-\lambda)^2 - 2b^2\right) = 0$$

$$\Rightarrow \lambda_1 = a; \qquad \lambda_{2,3} = a \pm b\sqrt{2}$$

$$\lambda_1 = a \Rightarrow \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix} \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ e_{1,3} \end{bmatrix} = a \begin{bmatrix} e_{1,1} \\ e_{1,2} \\ e_{1,3} \end{bmatrix} \Rightarrow \begin{matrix} ae_{1,1} + be_{1,2} = ae_{1,1} \\ be_{1,1} + ae_{1,2} + be_{1,3} = ae_{1,2} \\ be_{1,2} + ae_{1,3} = ae_{1,3} \end{matrix}$$

$$\Rightarrow \begin{matrix} e_{1,2} = 0 \\ e_{1,1} + e_{1,3} = 0 \end{matrix} \Rightarrow e_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

Example: 3 × 3 Symmetric Matrix

$$\lambda_{2,3} = a \pm b\sqrt{2} \quad \Rightarrow \quad \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix} \begin{bmatrix} e_{i,1} \\ e_{i,2} \\ e_{i,3} \end{bmatrix} = \left(a \pm b\sqrt{2}\right) \begin{bmatrix} e_{i,1} \\ e_{i,2} \\ e_{i,3} \end{bmatrix}; \quad i = 2,3$$

$$ae_{i,1} + be_{i,2} = \left(a \pm b\sqrt{2}\right) e_{i,1}$$

$$\Rightarrow \quad be_{i,1} + ae_{i,2} + be_{1,3} = \left(a \pm b\sqrt{2}\right) e_{i,2}$$

$$be_{i,2} + ae_{i,3} = \left(a \pm b\sqrt{2}\right) e_{i,3}$$

$$\Rightarrow \quad \begin{matrix} \left(e_{i,1} + e_{i,3}\right) = \pm e_{i,2}\sqrt{2} \\ e_{i,1} = e_{i,3}; \quad i = 2,3 \end{matrix} \quad \Rightarrow \quad e_{2,3} = \begin{bmatrix} 1/2 \\ \pm 1/\sqrt{2} \\ 1/2 \end{bmatrix}$$

Eigendecomposition

- **Matrix diagonalisation** (or **eigendecomposition**)
  - For a square $n \times n$ matrix with orthogonal eigenvectors
    - E.g. a symmetric real matrix
  - $E = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_n]$ - the orthonormal eigenvectors as the columns: $E_{ij} = e_{j,i}$
    - $AE = [\lambda_1\mathbf{e}_1\ \lambda_2\mathbf{e}_2\ \ldots\ \lambda_n\mathbf{e}_n]$ - the columns - the eigenvectors factored by their eigenvalues: $(AE)_{ij} = \lambda_j e_{j,i}$
    - The matrix E is orthogonal: $E^{\mathsf{T}}E = I$

Its transpose, E^{T}, is the (left) inverse, E^{-1}

- $E^{\mathsf{T}}AE = \Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ - the eigendecomposition $A = E\Lambda E^{\mathsf{T}}$

Eigendecomposition: Example

Eigenvectors $E = [\mathbf{e}_1 \quad \mathbf{e}_2]$

Eigenvalues $\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2\}$

Matrix to diagonalise

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{A} = \underbrace{\underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{E^{\mathsf{T}}}}_{\text{decomposition } E\Lambda E^{\mathsf{T}}}$$

$$\Rightarrow \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{E\Lambda} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{E^{\mathsf{T}}} = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{A}$$

Eigendecomposition: Examples

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{E^{\mathsf{T}}}$$

$$\underbrace{\begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} a & 0 & 0 \\ 0 & a+b\sqrt{2} & 0 \\ 0 & 0 & a-b\sqrt{2} \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix}}_{E^{\mathsf{T}}}$$

Singular Value Decomposition

- SVD is defined for an arbitrary rectangular matrix A: any $m \times n$ matrix A is decomposed as $A = UDV^{\mathsf{T}}$
  - U - the $m \times m$ matrix
    - Its columns are mutually orthogonal unit vectors
  - V - the $n \times n$ matrix
    - Its colums (i.e rows of V^{T}) are mutually orthogonal unit vectors
  - D - the $m \times n$ diagonal matrix: $\mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$
    - Singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$
  - U and V are not unique, but the singular values are fully determined by A

Properties of the SVD

- A square matrix A is non-singular if and only if all its singular values are different from zero
  - The values $\sigma_1, \ldots, \sigma_n$ show how close the matrix is to be singular
  - The **condition number** $C = \sigma_1 / \sigma_n$ gives the degree of singularity of A
  - If $1/C$ is comparable with the arithmetic precision of a computer, the matrix A is **ill-conditioned**
  - For all practical purposes, the ill-conditioned matrices should be considered singular
- If A is a rectangular matrix, the number of non-zero singular values equals the rank of A

Properties of the SVD

– **Effective rank** of A: the number of singular values greater than a fixed tolerance ε, being typically of order 10^{-6}

- **Inverse** of a square, non-singular matrix A:

$$A^{-1} = VD^{-1}U^{\mathsf{T}}$$

- **Pseudoinverse** of A, be A singular or not:

$$A^{+} = VD_0^{-1}U^{\mathsf{T}}$$

– D_0^{-1} is equal to D^{-1} for all nonzero singular values and zero otherwise

– If A is non-singular, then $D_0^{-1} = D^{-1}$ and $A^{+} = A^{-1}$

Properties of the SVD

- The columns of U corresponding to nonzero singular values span the range of A
  - The **range** of a $m \times n$ matrix A is the set of all m-vectors $A\mathbf{x}$, where $\mathbf{x}$ is any n-vector
- The columns of V corresponding to zero singular value span the null space of A
  - The **null space** of a $m \times n$ matrix A is the set of all n-vectors $\mathbf{x}$ such that $A\mathbf{x} = 0$
- The **Frobenius norm** of A:

$$\|A\|_{\mathrm{F}} = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij} = \sum_{j=1}^{n} \sigma_j^2 \Leftarrow \text{follows from the SVD } A = UDV^{\mathsf{T}}$$

Properties of the SVD

- The columns of U are the eigenvectors of AA^{T}
- The columns of V are the eigenvectors of $A^{\mathsf{T}}A$
- The squares of the nonzero singular values are the nonzero eigenvalues of both the nxn matrix $A^{\mathsf{T}}A$ and mxm matrix AA^{T}
- It holds that

$$A\mathbf{v}_i = \sigma_i\mathbf{u}_i \text{ and } A^{\mathsf{T}}\mathbf{u}_i = \sigma_i\mathbf{v}_i$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the columns of U and V corresponding to σ_i

Example: a Square 3×3 Matrix

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \Rightarrow \quad AA^{\mathsf{T}} = A^{\mathsf{T}}A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\Rightarrow$ Eigenvalues and eigenvectors for $AA^{\mathsf{T}} = A^{\mathsf{T}}A$:

$$\lambda_{1,2} = 1; \;\; \mathbf{e}_1 = \mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix}; \qquad \mathbf{e}_2 = \mathbf{u}_2 = \mathbf{v}_2 = \begin{bmatrix} 0 \\ -\sin\theta \\ \cos\theta \end{bmatrix};$$

$$\lambda_3 = 0; \;\; \mathbf{e}_3 = \mathbf{u}_3 = \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Example(cont): a Square 3×3 Matrix

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i; \quad i = 1,2,3 \quad \Rightarrow \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix} = \sigma_1 \begin{bmatrix} 0 \\ \cos\theta \\ \sin\theta \end{bmatrix} \quad \Rightarrow \sigma_1 = 1; \quad \cos\theta = \sin\theta = \frac{1}{\sqrt{2}}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \sigma_2 \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \Rightarrow \quad \sigma_2 = -1; \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \sigma_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \sigma_3 = 0 \quad \Rightarrow$$

$$A = \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix}}_{U} \underbrace{\text{diag}\{\sigma_1, \sigma_2, \sigma_3\}}_{D} \underbrace{\begin{bmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \mathbf{v}_3^\top \end{bmatrix}}_{V^\top} = \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix}}_{V^\top}$$

Example: a 3×2 Matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow AA^{\mathrm{T}} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$A^{\mathrm{T}}A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow$$

Eigenvalues and eigenvectors for AA^{T}: $\begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = 0 \quad \Rightarrow$

$$(1-\lambda)\left[(2-\lambda)(1-\lambda)-1\right]-(1-\lambda) = (3-\lambda)(1-\lambda)\lambda = 0 \quad \Rightarrow \lambda_1 = 3;\ \lambda_2 = 1;\ \lambda_3 = 0$$

$$\lambda_1 = 3;\ \mathbf{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}; \qquad \lambda_2 = 1;\ \mathbf{u}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}; \qquad \lambda_3 = 0;\ \mathbf{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Example (cont): a 3×2 Matrix

Eigenvalues and eigenvectors for $A^{\mathsf{T}}A$: $(2-\lambda)^2 - 1 = (3-\lambda)(1-\lambda) = 0 \;\Rightarrow\; \lambda_1 = 3;\; \lambda_2 = 1$

$$\lambda_1 = 3;\; \mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix};\; \lambda_2 = 1;\; \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}; \;\Rightarrow\; A\mathbf{v}_i = \sigma_i \mathbf{u}_i;\; i = 1,2 \;\Rightarrow$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \sigma_1 \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \Rightarrow \sigma_1 = \sqrt{3}; \qquad \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \sigma_2 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \Rightarrow \sigma_2 = -1$$

$$\Rightarrow \qquad A = \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix}}_{U} \underbrace{\text{diag}\{\sigma_1, \sigma_2\}}_{D} \underbrace{\begin{bmatrix} \mathbf{v}_1^{\mathsf{T}} \\ \mathbf{v}_2^{\mathsf{T}} \end{bmatrix}}_{V^{\mathsf{T}}} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{V^{\mathsf{T}}}$$

Computing the SVD…

- Another (equivalent) definition of SVD for computing: the $m \times n$ matrix U and $n \times n$ matrices D and V
  - Smaller memory space for U,D,V: $mn+2n^2$ rather than m^2+mn+n^2
- C-subroutine `int svd(int m, int n,`

```
double **A,
double *W,          // n non-negative SVs of A
int matu,           // 1 if the matrix U is desired
double **U,
int matv,           // 1 if the matrix V is desired
double **V,
double *rv1 )       // a working array of size n
```

Least Squares / Pseudoinverse

$A\mathbf{x} = \mathbf{b}$: an over-determined system of m linear equations

- A - an $m \times n$ matrix of coefficients
- $\mathbf{b}$ - an m-dimensional data vector
- $\mathbf{x}$ - the n-dimensional solution vector to be found
- If not all the components of $\mathbf{b}$ are null, the optimal in the least square sense (and the shortest-length) solution is:

$$\mathbf{x}^* = (A^{\mathsf{T}}A)^{+}A^{\mathsf{T}}\,\mathbf{b}$$

where B^{+} denotes the **pseudoinverse** of a matrix B

Least Squares Solution

Minimising the total squared error:

$$\|A\mathbf{x}-\mathbf{b}\|^2 = (A\mathbf{x}-\mathbf{b})^\top (A\mathbf{x}-\mathbf{b}) = (\mathbf{x}^\top A^\top - \mathbf{b}^\top)(A\mathbf{x}-\mathbf{b})$$

$$= \mathbf{x}^\top A^\top A\mathbf{x} - \mathbf{x}^\top A^\top \mathbf{b} - \mathbf{b}^\top A\mathbf{x} + \mathbf{b}^\top \mathbf{b} = \mathbf{x}^\top A^\top A\mathbf{x} - 2\mathbf{x}^\top A^\top \mathbf{b} + \mathbf{b}^\top \mathbf{b}$$

$$\frac{\partial \|A\mathbf{x}-\mathbf{b}\|^2}{\partial \mathbf{x}} = 2A^\top A\mathbf{x} - 2A^\top \mathbf{b} = 0 \quad \Rightarrow \quad \mathbf{x}^* = \underbrace{\left(A^\top A\right)^+}_{\text{pseudoinverse of } A^\top A} A^\top \mathbf{b} = \underbrace{\left(A^\top A\right)^{-1} A^\top}_{\text{pseudoinverse } A^+ \text{ of } A \text{ if the inverse } (A^\top A)^{-1} \text{ exists}} \mathbf{b}$$

Moore-Penrose inverse

Features of the pseudoinverse (B - a $m \times n$ matrix):

$$B^+ B = I_{n\times n}; \quad BB^+B = B; \quad B^+BB^+ = B^+$$

$$\left(B^+B\right)^\top = B^+B; \quad \left(BB^+\right)^\top = BB^+$$

Moore-Penrose Inverse

- If inverse of $(A^\top A)$ exists, the matrix $A^{+}= (A^\top A)^{-1}A^\top$ is the **pseudoinverse**, or **Moore-Penrose inverse**
  - The pseudoinverse exists for any $m \times n$ matrix; $m > n$
  - If $m>n$ (more equations than unknowns), the pseudoinverse $(A^\top A)^{+}$ is more likely coincide with the **inverse**, $(A^\top A)^{-1}$, of $A^\top A$
  - Still it is better to compute the pseudoinverse of $A^\top A$ through SVD to account for the condition number of this matrix
  - **SVD** (to get the pseudoinverse): if $A=UDV^\top$; $D=\mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$; $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge 0$, then $A^{+}=V(D^\top D)^{-1}D^\top U^\top$

Homogeneous Linear Systems

- $A\mathbf{x} = \mathbf{0}$ with $m \geq n - 1$ and $\mathrm{rank}(A) = n - 1$
  - Nontrivial solution $\mathbf{x}^*$ unique up to a scale factor: by SVD
    - It is proportional to the eigenvector corresponding to the only zero eigenvalue of $A^\mathsf{T}A$ (because all other eigenvalues are positive):

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^\mathsf{T} A\mathbf{x} = \mathbf{x}^\mathsf{T} A^\mathsf{T} A\mathbf{x} \rightarrow \min_{\mathbf{x}:\mathbf{x}^\mathsf{T}\mathbf{x}=1} \quad \Rightarrow \quad \min_{\mathbf{x}} \left\{ \mathrm{L}(\mathbf{x}) = \mathbf{x}^\mathsf{T} A^\mathsf{T} A\mathbf{x} - \lambda\left(\mathbf{x}^\mathsf{T}\mathbf{x} - 1\right)\right\}$$

$$\tfrac{d}{d\mathbf{x}}\mathrm{L}(\mathbf{x}) = 2A^\mathsf{T}A\mathbf{x} - 2\lambda\mathbf{x} = 0 \quad \Rightarrow \quad A^\mathsf{T}A\mathbf{x} = \lambda\mathbf{x}$$

  - λ is an eigenvalue of $A^\mathsf{T}A$; $\mathbf{x}^* = \mathbf{e}_\lambda$ is the eigenvector for λ
  - The solution $\mathrm{L}(\mathbf{e}_\lambda) = \lambda$; hence, the minimum is for $\lambda = 0$
  - Equivalently by SVD: the column of V for the only null SV of A