Similar Matrices and Similarity Transformation

- ***A* and $M^{-1}AM$ are said to be *similar* where *M* is any invertible matrix.**
- **$M^{-1}AM$ is called similarity transformation of *A* and vice versa**
- ***Two similar matrices share the same eigenvalues***
- **If $B = M^{-1}AM$ then *A* and *B* have the same eigenvalues and an eigenvector *x* of *A* corresponds to an eigenvector $M^{-1}x$ of *B***

***Proof*: since $A=MBM^{-1}$**

$$Ax = \lambda x \Rightarrow MBM^{-1}x = \lambda x \Rightarrow B\left(M^{-1}x\right) = \lambda\left(M^{-1}x\right)$$

Or we can look at the two determinants:

$$\det(B-\lambda I) = \det\left(M^{-1}AM - \lambda I\right) = \det\left(M^{-1}(A-\lambda I)M\right)$$
$$= \det M^{-1} \det(A-\lambda I) \det M = \det(A-\lambda I)$$

Example: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ **with eigenvalues 1 and 0**

If $M = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$ then $B = M^{-1}AM = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}$: triangular with eigenvalues 1 and 0

If $M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ then $B = M^{-1}AM = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$: projection with eigenvalues 1 and 0

If $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $B = M^{-1}AM =$ an arbitrary matrix with eigenvalues 1 and 0

Similarity Transformation = Change of Basis

Recall that

- **Every linear transformation is represented by a matrix**
- **The transformation matrix depends on the choice of basis**
- ***Similar matrices represent the same transformation with respect to different bases***
- **Let a transformation be *T* and a basis be $v_1, \ldots, v_n$. The *j*th column of the transformation *A* comes from applying *T* to v_j:**

  Tv_j=combination of the v's$=a_{1j}v_1+\cdots+a_{nj}v_n$

  Similarly, let the basis be $V_1, \ldots, V_n$ Then a new transformation matrix *B* can be constructed:

  TV_j=combination of the V's$=b_{1j}V_1+\cdots+b_{nj}V_n = \sum_{i=1}^{n} b_{ij}V_i$

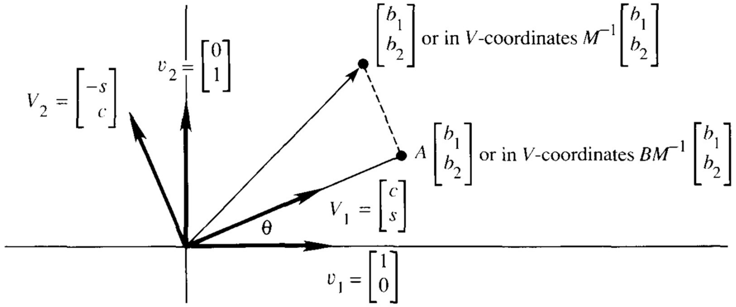
  In the mean time, $V_j = m_{1j}v_1+\cdots+m_{nj}v_n = \sum_{i=1}^{n} m_{ij}v_i$

- **The matrices *A* and *B* that represent the same linear transformation *T* with respect to two different bases *v* and *V* are *similar*:**

$$\begin{matrix} [T]_{V \text{ to } V} & = & [I]_{v \text{ to } V} & [T]_{v \text{ to } v} & [I]_{V \text{ to } v} \\ B & = & M^{-1} & A & M \end{matrix}$$

- **Transformation that changes only the basis is called *Identity Transformation*.**

Change of Basis Example – Projection



- **Transformation T is *projection* onto the line L at angle θ.**
- **Two bases: v_1=(1, 0) v_2=(0, 1) and V_1=(c, s) V_2=($-s$, c)**
- $B = [T]_{V \text{ to } V} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
- $M = [I]_{V \text{ to } v} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$
- $A = [T]_{v \text{ to } v} = MBM^{-1} = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}$
- **While *Elimination* preserves the *null space* and *row space*, *similarity transformation* preserves the *eigenvalues*.**

A and Λ are similar by changing the basis to the eigenvectors because $S^{-1}AS=\Lambda$

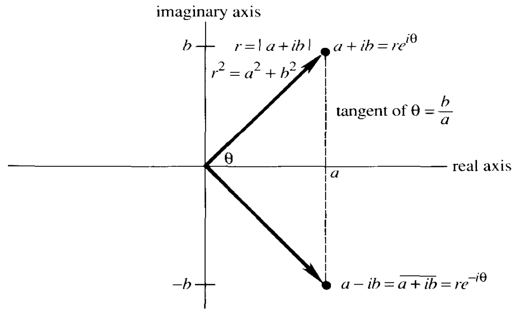
- **Numerically, eigenvalues are actually calculated by a sequence of simple similarities. The matrix gradually becomes triangular to find the eigenvalues on the diagonal (Chapter 7).**

Symmetric Matrix with Complex Numbers

- **Expressing u_0 as a linear combination of eigenvectors, how?**
- **What if the eigenvectors are orthogonal to one another?**
- **Our interest is symmetric matrices: Spectral Theorem**

  1. *symmetric matrix has <u>real eigenvalues</u>*

  2. *Its <u>eigenvectors can be chosen orthonormal</u>*
- **General case: "symmetric" (Hermitian) complex matrices**



Three properties of complex numbers and conjugates:

1. $\overline{(a+ib)(c+id)} = (ac-bd) - i(bc+ad) = \overline{(a+ib)}\,\overline{(c+id)}$

2. $\overline{(a+c)+i(b+d)} = (a+c) - i(b+d) = \overline{(a+ib)} + \overline{(c+id)}$

3. $(a+ib)(a-ib) = a^2 + b^2 = r^2$ **and r is called $|a+ib|$**

If $a = r\cos\theta$ and $b = r\sin\theta$, $a+ib = r(\cos\theta + i\sin\theta) = re^{i\theta}$

Example: x=3+4i $\Rightarrow x\bar{x} = (3+4i)(3-4i) = 25 = |x|^2 = r^2$ **and**

$$\frac{2+i}{3+4i} = \frac{2+i}{3+4i}\,\frac{3-4i}{3-4i} = \frac{10-5i}{25}$$

Complex Matrices: Lengths and Inner Products

- **Space C^n: all vectors x with n complex entries**

$$x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad x_j = a_j + ib_j$$

- **Complex vectors $v_1,\dots, v_k$ are dependent if there exists nontrivial ($c_1,\dots, c_k$) that produces $c_1v_1+\cdots c_kv_k$=0**
- **Length of complex vectors:**

$$\|x\|^2 = |x_1|^2 + \cdots + |x_n|^2$$

- **Inner product of two complex vectors:**

$$\bar{x}^{\mathrm{T}} y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$$

Example: x=(1+i, 3i) and y=(4, 2-i)

$\bar{x}^{\mathrm{T}} y = (1-i)4 + (-3i)(2-i) = 1-10i$

$\bar{x}^{\mathrm{T}} x = \overline{(1+i)}(1+i) + \overline{(3i)}(3i) = 2+9 = \|x\|^2$ **= (length of x)^2**

Hermitian Transpose and Hermitian Matrices

- **Definition: Hermitian Transpose**

$A^H = \bar{A}^{\mathrm{T}} = A^{\mathrm{H}}$ with entries $\left(A^{\mathrm{H}}\right)_{ij} = \overline{A_{ji}}$

Example: $\begin{bmatrix} 2+i & 3i \\ 4-i & 5 \\ 0 & 0 \end{bmatrix}^{\mathrm{H}} = \begin{bmatrix} 2-i & 4+i & 0 \\ -3i & 5 & 0 \end{bmatrix}$

- **Inner products of complex vectors x and y:** $\bar{x}^T y = x^H y$
- **Complex vectors x and y are orthogonal if $x^H y$=0**
- **Length of x=**$\|x\| = (x^H x)^{1/2}$
- **Since $(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$, $(AB)^{\mathrm{H}}=B^{\mathrm{H}}A^{\mathrm{H}}$**
- **Hermitian matrix: "symmetric" complex matrix**

$$\boxed{A=A^{\mathrm{H}}}$$

Example: $A=\begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = A^{\mathrm{H}}$ **that is,** $a_{ij} = \bar{a}_{ji}$

- **A real symmetric matrix is also *Hermitian***

⇒ A real symmetric matrix is a special case of Hermitian matrices

Three Properties of Hermitian Matrices

- **Property 1: If $A=A^H$, then $x^H Ax$ is real**

***Proof*: $(x^H Ax)^H = x^H A^H x^{HH} = x^H Ax$ only a real number's conjugate equals to itself.**

Example: $x^H Ax = \begin{bmatrix}\bar{u} & \bar{v}\end{bmatrix}\begin{bmatrix}2 & 3-3i \\ 3+3i & 5\end{bmatrix}\begin{bmatrix}u \\ v\end{bmatrix}$ **is real (?)**

$$= 2\bar{u}u + 5\bar{v}v + (3-3i)\bar{u}v + (3+3i)u\bar{v}$$

- **Property 2: Every eigenvalue of Hermitian matrix is real**

***Proof*: $Ax=\lambda x \Rightarrow x^H Ax = \lambda x^H x \Rightarrow x^H x$ real and positive $\Rightarrow \lambda$ real**

Example: $|A-\lambda I| = \begin{vmatrix}2-\lambda & 3-3i \\ 3+3i & 5-\lambda\end{vmatrix} = \lambda^2 - 7\lambda + 10 - |3-3i|^2$

$$= \lambda^2 - 7\lambda - 8 = (\lambda-8)(\lambda+1)$$

- **Property 3: The eigenvectors, from *different eigenvalues*, of a Hermitian matrix are orthogonal to one another.**

***Proof*: Let $Ax=\lambda_1 x$ and $Ay=\lambda_2 y$**

$$(\lambda_1 x)^H y = (Ax)^H y = x^H Ay = x^H(\lambda_2 y); \quad \because \lambda_1 \neq \lambda_2 \therefore x^H y = 0$$

Example:

$$(A-8I)x = \begin{bmatrix}-6 & 3-3i \\ 3+3i & -3\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix} \qquad x = \begin{bmatrix}1 \\ 1+i\end{bmatrix}$$

$$(A+I)y = \begin{bmatrix}3 & 3-3i \\ 3+3i & 6\end{bmatrix}\begin{bmatrix}y_1 \\ y_2\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix} \qquad y = \begin{bmatrix}1-i \\ -1\end{bmatrix}$$

And $x^H y = \begin{bmatrix}1 & 1-i\end{bmatrix}\begin{bmatrix}1-i \\ -1\end{bmatrix} = 0$

Diagonalization of Real Symmetric Matrices

- **Multiples of eigenvectors are eigenvectors $\Rightarrow$ orthogonal eigenvectors can be chosen to be orthonormal eigenvectors by choosing the multiplier $1/\|x\|$**

- **If $A=A^T$, the diagonalizing matrix S can be an orthogonal matrix Q**

- **A real symmetric matrix A can be factored into $A=Q\Lambda Q^T$ even with repeated eigenvalues (*to be proven later; we only prove the case with distinct eigenvalues*)**

$$A = Q\Lambda Q^T = \begin{bmatrix} & & \\ x_1 & \cdots & x_n \\ & & \end{bmatrix}\begin{bmatrix}\lambda_1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \lambda_n\end{bmatrix}\begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix}$$

$$= \lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T + \cdots + \lambda_n x_n x_n^T$$

$\Rightarrow$ a combination of one-dimensional <u>projections</u>: $x_i x_i^T$'s

Example: $A = \begin{bmatrix}2 & -1 \\ -1 & 2\end{bmatrix} = 3\begin{bmatrix}\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix} + \begin{bmatrix}\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$ **with**

$\begin{bmatrix}\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix}$ **to project onto** $x_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 \\ -1\end{bmatrix}$ **and** $\begin{bmatrix}\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\end{bmatrix}$ **to project onto** $x_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 \\ 1\end{bmatrix}$

Unitary Matrices – Complex Orthogonal Matrices

- ***U* is called a *Unitary Matrix* if**

$$\boxed{U^H U = I}$$

- $U^H U = I \Rightarrow UU^H = I$ **or** $U^H = U^{-1}$

Three properties of unitary matrices

1. $(Ux)^H(Uy) = x^H U^H U y = x^H y$ **and (by choosing** $y=x$**) lengths are preserved:** $\|Ux\|^2 = \|x\|^2$

2. Every eigenvalue of *U* has absolute value $|\lambda|=1$

$$\because Ux = \lambda x \Rightarrow \|Ux\| = \|x\| \Rightarrow \|Ux\| = |\lambda|\,\|x\| \therefore |\lambda| = 1$$

3. Eigenvectors from different eigenvalues are orthogonal

***Proof*: Let** $Ux=\lambda_1 x$ **and** $Uy=\lambda_2 y$

$$\lambda_1 x^H y = \lambda_1 (Ux)^H (Uy) = \lambda_1 (\lambda_1 x)^H (\lambda_2 y) = \lambda_1 \bar{\lambda}_1 \lambda_2 x^H y \Rightarrow$$

$$\lambda_1 x^H y = \lambda_2 x^H y \; \bar{\lambda}_1 \lambda_2 = 1 \quad \textbf{since} \quad \lambda_1 \bar{\lambda}_1 = 1 \quad \textbf{and}\; \lambda_1 \neq \lambda_2 \quad \Rightarrow x^H y = 0$$

Example: $U = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ **with eigenvalues** e^{it} **and** e^{-it}

$$\begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \begin{bmatrix} e^{it} & \\ & e^{-it} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}$$

Unitary Matrices and Skew-Hermitian Matrices

- **There exists a unitary *U* such that** $U^{-1}AU = U^H AU = \Lambda$ **where *A* is Hermitian and** Λ **is a matrix with real eigenvalues** λ_i**'s as its diagonal entries.**
- **Skew-Hermitian Matrices:**

$$\boxed{K^H = -K}$$

- **If *A* is Hermitian then** $K=iA$ **is skew-Hermitian**

Example: $K = iA = \begin{bmatrix} 2i & 3+3i \\ -3+3i & 5i \end{bmatrix} = -K^H$

- **Three properties of skew-Hermitian matrices**
  1. **For any *x*,** $x^H Kx$ **is pure imaginary**
  2. **Every eigenvalue of *K* is pure imaginary**
  3. **Eigenvectors corresponding to different eigenvalues are orthogonal**
- **There is a unitary *U* such that** $U^{-1}KU = \Lambda$

Real vs. Complex

R^n **= space of vectors with n components** $\leftrightarrow$ **C**^{n} **= space of vectors with n complex components**

Length: $\|x\|^2 = x_1^2 + \cdots + x_n^2$ $\leftrightarrow$ $\|x\|^2 = |x_1|^2 + \cdots + |x_n|^2$

Transpose: $A_{ij}^T = A_{ji}$ $\leftrightarrow$ **Hermitian transpose:** $A_{ij}^H = \overline{A_{ji}}$

$(AB)^T = B^T A^T$ $\leftrightarrow$ $(AB)^H = B^H A^H$

Inner product:

$x^T y = x_1 y_1 + \cdots + x_n y_n$ $\leftrightarrow$ $x^H y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$

$(Ax)^T y = x^T (A^T y)$ $\leftrightarrow$ $(Ax)^H y = x^H (A^H y)$

Orthogonality: $x^T y = 0$ $\leftrightarrow$ $x^H y = 0$

Symmetric matrices: $A^T = A$ $\leftrightarrow$ **Hermitian matrices:** $A^H = A$

$A = Q\Lambda Q^{-1} = Q\Lambda Q^T (\text{real } \Lambda)$ $\leftrightarrow$ $A = U\Lambda U^{-1} = U\Lambda U^H (\text{real } \Lambda)$

Skew-symmetric: $K^T = -K$ $\leftrightarrow$ **Skew-Hermitian:** $K^H = -K$

Schur's Lemma – Trianglizing by a Unitary M

- **For any matrix A, there is a unitary matrix $M=U$ such that $U^{-1}AU=T$ is upper triangular. The eigenvalues of A, shared by the similar matrix T, appear along its diagonal.**

***Proof*: Take 4 by 4 matrix A as an example. There is at least one eigenvalue λ_1. Unitary matrix U_1 can be constructed by the corresponding eigenvector x_1 as the first vector and by Gram-Schmidt process for the subsequent 3 vectors. Then,**

$$AU_1 = U_1\begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \text{ or } U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

Second step: look at the lower right 3 by 3 matrix and there exists at least one eigenvalue λ_2 so that its corresponding eigenvector and two perpendicular vectors (Gram-Schmidt) can form M_2 and

$$U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix} \quad \text{such that} \quad U_2^{-1}\left(U_1^{-1}AU_1\right)U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

Last step: $$U_3^{-1}\left(U_2^{-1}U_1^{-1}AU_1U_2\right)U_3 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & * \end{bmatrix} = T$$

- **The product $U=U_1U_2U_3$ is still unitary (exercise)**

Schur's Lemma and Spectral Theorem

- **Schur's lemma can be used to check the stability without diagonalizing A:**

  1. The powers A^k approach zero when all $|\lambda_i|<1$

  2. The exponentials e^{At} approach zero when all Re $\lambda_i<0$

Example: $A=\begin{bmatrix}2 & -1\\ 1 & 0\end{bmatrix}$ **has the eigenvalue λ=1 (twice)**

$$U^{-1}AU=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}\begin{bmatrix}2 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}$$

- **(Spectral theorem) Every real symmetric matrix can be diagonalized by an orthogonal matrix and every Hermitian matrix can be diagonalized by a unitary matrix: (real case) $Q^TAQ=\Lambda$ (complex case) $U^HAU=\Lambda$**

  The columns of Q (or U) contain a complete set of orthonormal eigenvectors

***Proof*:**

1. If A is Hermitian then so is $U^{-1}AU$:

$$\left(U^{-1}AU\right)^{H}=U^{H}A^{H}\left(U^{-1}\right)^{H}=U^{-1}AU$$

2. If a symmetric or Hermitian matrix is also triangular, it must be diagonal

Example of Spectral Theorem

$$A=\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}$$

Repeated eigenvalues $\lambda_1=\lambda_2=1$ and $\lambda_3=-1$ with eigenvectors (also columns of an orthogonal matrix Q and $A=Q\Lambda Q^T$)

$$x_1=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}\quad x_2=\begin{bmatrix}0\\0\\1\end{bmatrix}\quad x_3=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix}$$

***A* becomes:**

$$A=\sum\lambda_i x_i x_i^T=\lambda_1\begin{bmatrix}\frac12 & \frac12 & 0\\ \frac12 & \frac12 & 0\\ 0 & 0 & 0\end{bmatrix}+\lambda_2\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}+\lambda_3\begin{bmatrix}\frac12 & -\frac12 & 0\\ -\frac12 & \frac12 & 0\\ 0 & 0 & 0\end{bmatrix}$$

$\because \lambda_1=\lambda_2$ **the first two projections can be combined:**

$$\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}=\lambda_1P_1+\lambda_3P_3=(+1)\begin{bmatrix}\frac12 & \frac12 & 0\\ \frac12 & \frac12 & 0\\ 0 & 0 & 1\end{bmatrix}+(-1)\begin{bmatrix}\frac12 & -\frac12 & 0\\ -\frac12 & \frac12 & 0\\ 0 & 0 & 0\end{bmatrix}$$

P_1: projection onto the plane spanned by x_1 and x_2

P_2: projection onto the line with the same direction as x_3

- ***Every Hermitian matrix with k different eigenvalues has its own "spectral decomposition" into*** $A=\lambda_1P_1+\cdots\lambda_kP_k$***, where P_i is the projection onto the eigensapce for λ_i.***

Diagonalizing Normal Matrices

- The matrix N is *normal* if it commutes with N^H: $NN^H=N^HN$.
- For normal matrices, and no others, the triangular $T=U^{-1}NU$ is the diagonal matrix Λ. Normal matrices are exactly those that possess a complete set of orthonormal eigenvectors.

Proof:

1. If N is normal then so is $T=U^{-1}NU$

$TT^H=U^{-1}NUU^HN^HU=U^{-1}NN^HU=U^{-1}N^HNU=U^HN^HUU^{-1}NU=T^HT$

2. A triangular T that is normal must be diagonal (exercises 5.6.19-20)

Diagonalizing General Matrices – Jordan Form

- Goal: make $M^{-1}AM$ as nearly diagonal as possible
- If A has s independent eigenvectors, it is similar to a matrix with s blocks:

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & J_s \end{bmatrix}$$

Each Jordan block J_i is a triangular matrix with only a single eigenvalue λ_i and only one eigenvector.

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}$$

- When the block has order m>1, the eigenvalue λ_i is repeated m times and there are m-1 1's above the diagonal.
- The same eigenvalue λ_i may appear in several blocks, if it corresponds to several independent eigenvectors.
- Two matrices, shared the same Jordan form, are similar

Jordan Form - Examples

Example 1: $T=\begin{bmatrix}1&2\\0&1\end{bmatrix}$ **and** $A=\begin{bmatrix}2&-1\\1&0\end{bmatrix}$ **and** $B=\begin{bmatrix}1&0\\1&1\end{bmatrix}$ **share** $J=\begin{bmatrix}1&1\\0&1\end{bmatrix}$

$$M^{-1}TM=\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&2\\0&1\end{bmatrix}\begin{bmatrix}1&0\\0&\frac{1}{2}\end{bmatrix}=\begin{bmatrix}1&1\\0&1\end{bmatrix}=J \quad \textbf{(2}\rightarrow\textbf{1)}$$

$U^{-1}AU=\begin{bmatrix}1&2\\0&1\end{bmatrix}=T$ and then $M^{-1}TM=\begin{bmatrix}1&1\\0&1\end{bmatrix}=J$ **(make triangular and then 2→1)**

$$P^{-1}BP=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}1&1\\0&1\end{bmatrix}=J \quad \textbf{(permutations)}$$

Example 2: $A=\begin{bmatrix}0&1&2\\0&0&1\\0&0&0\end{bmatrix}$ and $B=\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}$

Eigenvalue 0 repeats 3 times ⇒ three possibilities:

1. a single 3×3 block; 2. one 2×2 and one 1×1; 3. three 1×1

$$J_1=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}\quad J_2=\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix}\quad\text{and}\quad J_3=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

For A, there is only one eigenvector (1, 0, 0): J_1

For B, there is an additional eigenvector (0, 1, 0): J_2

Applications of Jordan Form to Difference and Differential Equations

- **For solution to a difference equation:**

$$A^k=\left(MJM^{-1}\right)\left(MJM^{-1}\right)\cdots\left(MJM^{-1}\right)=MJ^kM^{-1} \quad \textbf{and}$$

$$J_i^n=\begin{bmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{bmatrix}^n=\begin{bmatrix}\lambda^n&n\lambda^{n-1}&n(n-1)\lambda^{n-2}\\0&\lambda^n&n\lambda^{n-1}\\0&0&\lambda^n\end{bmatrix}$$

- **For solution to a differential equation:**

$$e^{J_it}=\begin{bmatrix}e^{\lambda t}&te^{\lambda t}&\frac{1}{2}t^2e^{\lambda t}\\0&e^{\lambda t}&te^{\lambda t}\\0&0&e^{\lambda t}\end{bmatrix}$$

Remember: $I+J_it+(J_it)^2/2!+\cdots$ **⇒ e.g.** $1+\lambda t+(\lambda t)^2/2!+\cdots=e^{\lambda t}$ **on the diagonal.**

Example: $\begin{bmatrix}du_1/dt\\du_2/dt\\du_3/dt\end{bmatrix}=\begin{bmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{bmatrix}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}$ starting from $u_0=\begin{bmatrix}0\\0\\1\end{bmatrix}$

Back-substitution: $u_3=e^{\lambda t}$ **then** $du_2/dt=\lambda u_2+u_3 \Rightarrow u_2=te^{\lambda t}$

$du_1/dt=\lambda u_1+u_2 \Rightarrow u_3=t^2e^{\lambda t}/2$

- **General solutions:**

$$\text{if}\quad u_{k+1}=Au_k\quad\text{then}\quad u_k=A^ku_0=MJ^kM^{-1}u_0$$
$$\text{if}\quad du/dt=Au\quad\text{then}\quad u=e^{At}u_0=Me^{Jt}M^{-1}u_0$$

Quadratic Function: $f=ax^2+2bxy+cy^2$

What do we know about $f=ax^2+2bxy+cy^2$?

- **f=0 at the point (0, 0) ⇒ stationary point**
- **Stationary point: the point at which $\partial f/\partial x$=0 and $\partial f/\partial y$=0**

$$f = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

- **f is strictly positive (negative) except at (0, 0) if a>0 (a<0) and $ac>b^2$ ($c>b^2/a$ if a>0 or $c< b^2/a$ if a<0):**

  Algebraically: *Positive* (*Negative*) *Definite*

  Geometrically: a (upside down) bowl

- **f is nonnegative (none-positive) if $a\geq(\leq)0$ and $ac\geq b^2$:**

  Algebraically: *Positive* (*Negative*) *Semidefinite*

  Geometrically: a (upside down) trough, degenerated from a (upside down) bowl

Are there other possibilities?

Indefinite and Saddle Point

- **f is ambiguous if $ac-b^2$ is negative**

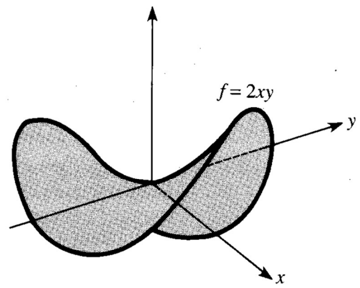
  Algebraically: *Indefinite*

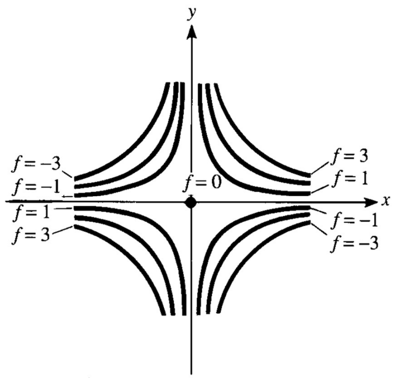
  Geometrically: a saddle

Example: $f_1 = 2xy$ and $f_2 = x^2 - y^2 \Rightarrow$ **$ac-b^2$ = -1**

f_2: the stationary point is minimum along x-axis and is the maximum along the y-axis ⇒ saddle point

f_1: turn f_2 45°





Example: $f(x, y) = 2x^2 + 4xy + y^2$

Quadratic Function and General Function

- **The quadratic form is positive definite if and only if *a*>0 and *ac*-*b*²>0. Correspondingly, F has a (local) minimum at *x*=*y*=0 if and only if its *first derivatives are zero* and**

$$\frac{\partial^2 F}{\partial x^2}(0{,}0)>0 \qquad \left[\frac{\partial^2 F}{\partial x^2}(0{,}0)\right]\left[\frac{\partial^2 F}{\partial y^2}(0{,}0)\right]>\left[\frac{\partial^2 F}{\partial x\,\partial y}(0{,}0)\right]^2$$

Example: $F(x,y)=7+2(x+y)^2-y\sin y-x^3$

$$f(x,y)=2x^2+4xy+y^2$$

First derivative:

$$\frac{\partial F}{\partial x}=4(x+y)-3x^2=0 \quad \text{and} \quad \frac{\partial F}{\partial y}=4(x+y)-y\cos y-\sin y=0$$

$$\frac{\partial f}{\partial x}=4x+4y=0 \quad \text{and} \quad \frac{\partial f}{\partial y}=4x+2y=0$$

Second derivative:

$$\frac{\partial^2 F}{\partial x^2}=4-6x=4=2a \qquad \frac{\partial^2 f}{\partial x^2}=4=2a$$

$$\frac{\partial^2 F}{\partial x\,\partial y}=\frac{\partial^2 F}{\partial y\,\partial x}=4=2b \qquad \frac{\partial^2 f}{\partial x\,\partial y}=\frac{\partial^2 f}{\partial y\,\partial x}=4=2b$$

$$\frac{\partial^2 F}{\partial y^2}=4+y\sin y-2\cos y=2=2c \qquad \frac{\partial^2 f}{\partial y^2}=2=2c$$

For stationary point at (α, β): Check

$$f(x,y)=\underbrace{\frac{1}{2}\frac{\partial^2 F}{\partial x^2}(\alpha,\beta)}_{a}x^2+\underbrace{\frac{\partial^2 F}{\partial x\,\partial y}(\alpha,\beta)}_{2b}xy+\underbrace{\frac{1}{2}\frac{\partial^2 F}{\partial y^2}(\alpha,\beta)}_{c}y^2$$

Quadratic Function in Matrix Form

$$ax^2+2bxy+cy^2=\begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}a & b\\ b & c\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$$

- **Power of linear algebra (or mathematics): extension to something hard to envision:**

$$x^{\mathrm{T}}Ax=\begin{bmatrix}x_1 & x_2 & \cdot & x_n\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & \cdot & a_{1n}\\ a_{21} & a_{22} & \cdot & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{n1} & a_{n2} & \cdot & a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \cdot\\ x_n\end{bmatrix}$$

$$\Rightarrow \quad =a_{11}x_1^2+a_{12}x_1x_2+a_{21}x_2x_1+\cdots+a_{nn}x_n^2$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

⇒ ***f* = *x*ᵀ*Ax* where *A* is a symmetric matrix**

Example: $f=2x_1^2-2x_1x_2+2x_2^2-2x_2x_3+2x_3^2$

$$A=\begin{bmatrix}2 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}\to \text{minimum}$$

- **Matrix *A* with entries:** $a_{ij}=\dfrac{\partial F}{\partial x_i \partial x_j}$ **(Hessian Matrix) for any function $F(x_1,\ldots,x_n)$ ⇒ *F* has a minimum when the pure quadratic $f=x^{\mathrm{T}}Ax$ is positive. Think of Taylor series:**

$$F(x)=F(0)+x^{\mathrm{T}}(\text{grad }F)+\tfrac{1}{2}x^{\mathrm{T}}Ax+\text{3rd-order terms}$$

where grad F =$(\partial F/\partial x_1,\ldots,\partial F/\partial x_n)$

Test for Positive Definiteness

if $A=\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ we need $a>0$ $ac-b^2>0$

1. both eigenvalues are positive

2. det [*a*]>0 and det *A*>0

$$ax^2+2bxy+cy^2=a\left(x+\frac{b}{a}y\right)^2+\frac{ac-b^2}{a}y^2$$

3. pivots *a*>0 and (*ac-b*^{2})/*a*>0

- **Tests for a real symmetric matrix *A* to be positive definite:**

(I) x^TAx>0 for all nonzero vectors *x* (definition)

(II) All the eigenvalues of *A*: λ_i>0

(III) All the upper left submatrices A_k have positive determinants

Remember:

$$A_1=[a_{11}] \quad A_2=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad A_3=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdots \quad A_n=A$$

(IV) All the pivots d_i>0

(V) There exists a matrix *R* with independent columns such that

$A=R^TR$

Proof: Positive Definiteness Tests

- **(I)→(II): *A* symmetric**

Choose unit eigenvectors x_i's with corresponding eigenvalues λ_i's

then $Ax_i=\lambda_i x_i$ so $x_i^TAx_i=x_i^T\lambda_i x_i=\lambda_i$**>0**

- **(II)→(I): all λ_i>0**

***A* is real symmetric ⇒ orthonormal eigenvectors**

Any *x* can be written as a combination $c_1x_1+\ldots+c_nx_n$

$$Ax=c_1Ax_1+\cdots+c_nAx_n=c_1\lambda_1x_1+\cdots+c_n\lambda_nx_n$$

$$\Rightarrow \begin{aligned} x^TAx &= \left(c_1x_1^T+\cdots+c_nx_n^T\right)\left(c_1\lambda_1x_1+\cdots+c_n\lambda_nx_n\right) \\ &= c_1^2\lambda_1+\cdots+c_n^2\lambda_n>0 \end{aligned}$$

- **(I)→(III):**

$$\det A=\lambda_1\lambda_2\cdots\lambda_n>0$$

$$x^TAx=\begin{bmatrix} x_k^T & 0 \end{bmatrix}\begin{bmatrix} A_k & * \\ * & * \end{bmatrix}\begin{bmatrix} x_k \\ 0 \end{bmatrix}=x_k^TA_kx_k\mathbf{>0}$$

- **(III)→(IV):** $d_k=\dfrac{\det A_k}{\det A_{k-1}}$**>0**

- **(IV)→(I): *A* is symmetric ⇒ *A=LDL*^{T}**

  ⇒ pivots are coefficients of square terms

- **(V): $x^TAx=x^TR^TRx=||Rx||^2$**

Ellipsoids in *n* Dimensions

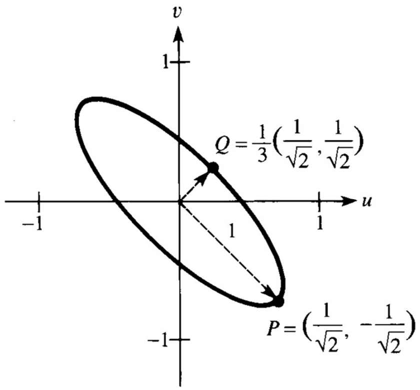
Ellipse and Ellipsoids: $x^{\mathrm{T}}Ax$=1

Example:

for $A=\begin{bmatrix}4 & & \\ & 1 & \\ & & \frac{1}{9}\end{bmatrix}$ the equation is $x^{\mathrm{T}}Ax=4x_1^2+x_2^2+\frac{1}{9}x_3^2=1$

Example: $A=\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}$ **and $x^{\mathrm{T}}Ax=5u^2+8uv+5v^2=1$**

$$5u^2+8uv+v^2=\left(\frac{u}{\sqrt{2}}-\frac{v}{\sqrt{2}}\right)^2+9\left(\frac{u}{\sqrt{2}}+\frac{v}{\sqrt{2}}\right)^2=1$$



Eigenvalues/vectors: 1, $(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})$ **and 9,** $(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$

Major axis: direction $(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})$**; length 1**

Minor axis: direction $(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$**; length** $1/\sqrt{9}=1/3$

Ellipsoids and Eigenvalues/Eigenvectors

$$x^{\mathrm{T}}Ax=\left(x^{\mathrm{T}}Q\right)\Lambda\left(Q^{\mathrm{T}}x\right)=y^{\mathrm{T}}\Lambda y=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2$$

- **$x\to y=Q^{\mathrm{T}}x$: rotate the axes to match axes of the ellipsoid**

Example:

$x^{\mathrm{T}}Ax=5u^2+8uv+5v^2=1\Rightarrow$ $A=\begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}$ **eigenvalues: 1, 9**

$\Rightarrow\ y_1^2+9y_2^2=1$

- **Suppose A is positive definite: $A=Q\Lambda Q^{\mathrm{T}}$ with $\lambda_i>0$. Then the rotation $y=Q^{\mathrm{T}}x$ simplifies $x^{\mathrm{T}}Ax=1$ to**

  $x^{\mathrm{T}}Q\Lambda Q^{\mathrm{T}}x=1$ or $y^{\mathrm{T}}\Lambda y=1$ or $\lambda_1 y_1^2+\cdots+\lambda_n y_n^2=1$

  This is the equation of an ellipsoid. Its axes have lengths $1/\sqrt{\lambda_1},\ldots,1/\sqrt{\lambda_n}$ **from the center. And in the original *x*-space they point along the eigenvectors.**

Semidefinite Tests

- **For A to be positive semidefinite**

(I') $x^{T}Ax \geq 0$ for all nonzero vectors x (definition)

(II') All the eigenvalues of A: $\lambda_i \geq 0$

(III') All principal submatrices have nonnegative determinants

(IV') All the pivots $d_i \geq 0$

(V') There exists a matrix R, possibly with dependent columns such that $A=R^{T}R$

Example:

$$A=\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

(I') $x^{T}Ax=(x_1-x_2)^2+(x_1-x_3)^2+(x_2-x_3)^2 \geq 0$

(II') eigenvalues: 0, 3, 3

(III') determinats of principal submatrices: 2 (1×1), 3(2×2), 0

(IV') $\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{2} & \frac{3}{2} \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$