Symmetric matrices and quadratic forms

- eigenvectors of symmetric matrices
- quadratic forms
- inequalities for quadratic forms
- positive semidefinite matrices

Eigenvalues & Eigenvectors

- char. polynomial: $\chi_A(\lambda) = \det(\lambda I - A)$ $\quad A \in \mathbb{R}^{n\times n}$

$\chi_A(\lambda) = 0 \longrightarrow \lambda_1, \dots, \lambda_n$: eigenvalues

$\lambda_i \in \mathbb{C} \to$ if $\chi_A(\lambda_0) = 0$

$\Rightarrow \chi_A(\bar{\lambda}_0) = 0$

- Diagonalizability: $\Leftrightarrow$ having n indep. eigenvectors : $v_1, \dots, v_n$

$\downarrow$

$T = [v_1, \dots, v_n]$

$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$$A = T \Lambda T^{-1}$$

- Symmetric (real) matrices:

① $\lambda_i \in \mathbb{R}$

② There exists an orthonormal set of eig. vectors

$Q = [v_1 \dots v_n] \to Q^{-1} = Q^T$

$\Rightarrow \boxed{A = Q \Lambda Q^T}$

③ $\lambda_i \neq \lambda_j \Rightarrow v_i^T v_j -$

Eigenvalues of symmetric matrices

1. if $A \in \mathbb{R}^{n\times n}$ is symmetric, *i.e.*, $A = A^\mathsf{T}$, then the eigenvalues of A are real

to see this, suppose $Av = \lambda v$, $v \neq 0$, $v \in \mathbb{C}^n$, then

$$\underbrace{\overline{v}^\mathsf{T}}_{1\times n}\underbrace{A}_{n\times n}\underbrace{v}_{n\times 1} = \overline{v}^\mathsf{T}(Av) = \lambda \overline{v}^\mathsf{T} v = \boxed{\lambda \sum_{i=1}^n |v_i|^2}$$

but also

$$\overline{v}^\mathsf{T} A v = \overline{(Av)}^\mathsf{T} v = \overline{(\lambda v)}^\mathsf{T} v = \boxed{\overline{\lambda} \sum_{i=1}^n |v_i|^2}$$

so we have $\lambda = \overline{\lambda}$, *i.e.*, $\lambda \in \mathbb{R}$ (hence, can assume $v \in \mathbb{R}^n$)

$a \underbrace{+bi}_{0} = a \underbrace{-bi}_{0}$

Eigenvectors of symmetric matrices

$$\boxed{uu^T} \quad n\times 1 \quad 1\times n \quad u \in \mathbb{R}^n$$

$$\boxed{y = Ax}$$

$$\begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} q_1^T \\ \vdots \\ q_n^T \end{bmatrix} = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots + \lambda_n q_n q_n^T$$

2. there is a set of n orthonormal eigenvectors of A

$\mathrm{rank}(AB) \le \mathrm{rank}(A) = 1$

$$\left.\begin{aligned} q_i^T q_i &= \|q_i\|^2 = 1 \\ q_i^T q_j &= 0 \end{aligned}\right\} \to Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}$$

$$\Rightarrow Q^T Q = Q Q^T = I$$
$$Q^{-1} = Q^T$$

- i.e., $q_1, \dots, q_n$ s.t. $Aq_i = \lambda_i q_i$, $\boxed{q_i^T q_j = \delta_{ij}}$
- in matrix form: there is an orthogonal Q s.t.

$$\underbrace{Q^{-1} A Q}_{\text{all diagonalizable matrices}} = \underline{Q^T A Q} = \Lambda$$

- hence we can express A as

$$A = Q \Lambda Q^T = \sum_{i=1}^{n} \underbrace{\lambda_i q_i q_i^T}_{\text{projection matrix}}$$

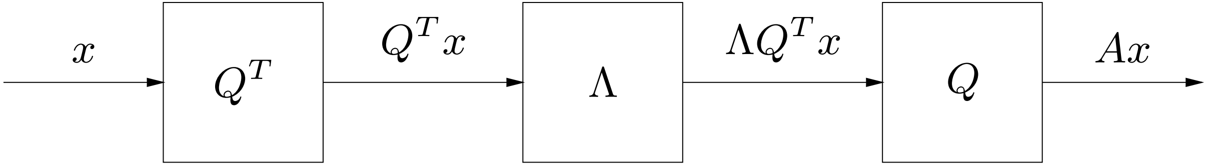
- in particular, q_i are both left and right eigenvectors

Interpretations

$$y = Ax = Q\Lambda Q^T x$$

$A = Q\Lambda Q^{\mathsf{T}}$ corresponds to

$$= Q(\Lambda(Q^T x))$$

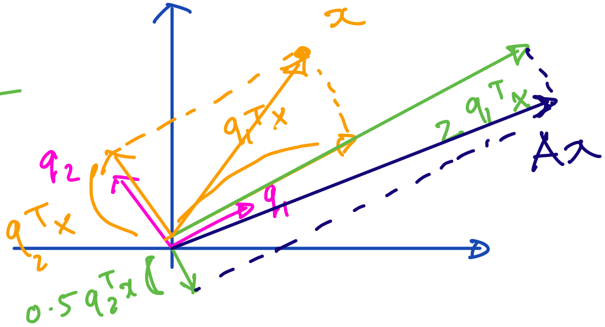


linear mapping $y = Ax$ can be decomposed as

- resolve into q_i coordinates
- scale coordinates by λ_i
- reconstitute with basis q_i

$\lambda_1 = 2$

$\lambda_2 = -0.5$



Geometrical interpretation

multiplication by A is the same as

- rotate by Q^T
- diagonal real scale ('dilation') by Λ
- rotate back by Q

decomposition

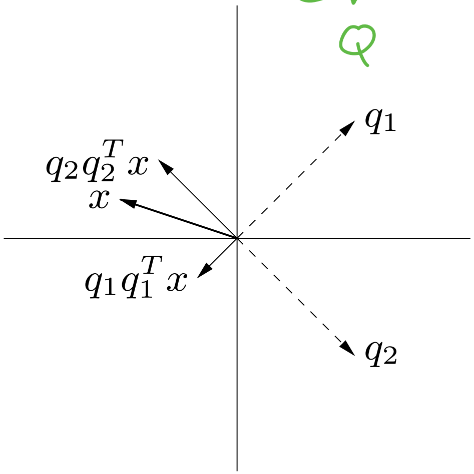
$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^\mathsf{T}$$

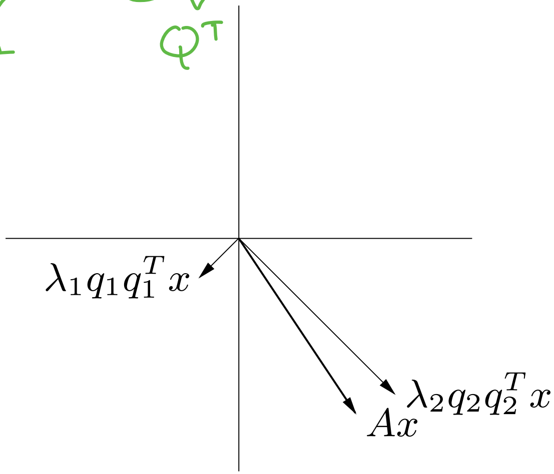
expresses A as linear combination of 1-dimensional projections

Example:

$$A = \begin{bmatrix} -1/2 & 3/2 \\ 3/2 & -1/2 \end{bmatrix}$$

$$= \underbrace{\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)}_{Q} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}}_{\Lambda} \underbrace{\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)^T}_{Q^T}$$





Proof

1. $\lambda_i \in \mathbb{R}$

2. There exists a set of eig vectors that is orthonormal ✓ ← I

3. eigenvectors corresponding to distinct eigenvalues are orthogonal X

- since λ_i distinct, have $v_1, \ldots, v_n$, a set of linearly independent eigenvectors of A

$$Av_i = \lambda_i v_i, \qquad ||v_i|| = 1$$

$n \times n$ matrix

sym. distinct eig. values.

$\lambda_j v_i^T v_j$

- then $v_i^\mathsf{T}(Av_j) = \lambda_j v_i^\mathsf{T} v_j = (Av_i)^\mathsf{T} v_j = \lambda_i v_i^\mathsf{T} v_j$

$v_i^T A^T = v_i^T A$

- and $(\lambda_i - \lambda_j) v_i^\mathsf{T} v_j = 0$
- for $i \neq j$, $\lambda_i \neq \lambda_j$, hence $v_i^\mathsf{T} v_j = 0$
- in this case we can say: eigenvectors *are* orthogonal
- in general case (λ_i not distinct) we must say: eigenvectors *can be chosen* to be orthogonal

Quadratic forms

generalization of ax^2 to $\mathbb{R}^n$

a *quadratic form* is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$(x^T A x)^T = x^T A x$$

$$x^T A^T x = -x^T A x \quad \text{if } A^T = -A$$

$$f(x) = \underbrace{x^T A x}_{1\times n \,\cdot\, n\times n \,\cdot\, n\times 1} = \sum_{i,j=1}^{n} A_{ij} x_i x_j$$

$A_{ii} x_i^2$

- in a quadratic form we may as well assume $A = A^T$ since

$$x^T A x = x^T ((A + A^T)/2) x$$

$((A + A^T)/2$ is called the *symmetric part* of A)

$$A = A_{sym} + A_{asym}$$

$$A_{sym}^T = A_{sym}$$

$$A_{asym}^T = -A_{asym}$$

$$A_{sym} = \frac{A + A^T}{2}, \quad A_{asym} = \frac{A - A^T}{2}$$

$$x^T A x = x^T A_{sym} x + \underbrace{x^T A_{asym} x}_{0}$$

$$x^T A x = x^T A_{sym} x$$

- **uniqueness:** if $x^T A x = x^T B x$ for all $x \in \mathbb{R}^n$ and $A = A^T$, $B = B^T$, then $A = B$

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \rightarrow x^T A x = \sum A_{ij} x_i x_j$$

$$= x_1^2 + \underbrace{2x_1x_2 + 4x_2x_1}_{6x_1x_2} + 3x_2^2$$

$$= \underbrace{x_1^2 + 3x_1x_2 + 3x_2x_1 + 3x_2^2}_{\begin{bmatrix} 1 & 3 \\ 3 & 3 \end{bmatrix}}$$

$$x^T A x = x^T B x \Rightarrow A = B$$

$$A^T = A, \quad B = B^T$$

$$x^T (A - B) x = 0 \rightarrow x = e_j \rightarrow x^T A x = A_{jj} = 0$$

$$x = \begin{bmatrix} 1 \\ 1 \\ 0 \\ \vdots \end{bmatrix} \rightarrow x^T A x = \underbrace{A_{11}}_{0} + \underbrace{A_{22}}_{0} + 2A_{12} = 0 \rightarrow A_{12} = 0$$

$$f(e_i) = \boxed{A_{ii} = 0}$$

$$f(e_i + e_j) = \underbrace{A_{ii}}_{0} + \underbrace{A_{jj}}_{0} + 2A_{ij} = 0 \rightarrow \boxed{A_{ij} = 0}$$

Examples

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

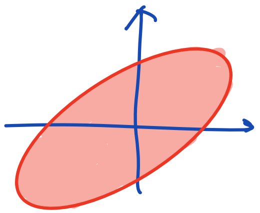
quadratic forms $B \in \mathbb{R}^{m \times n}$

- $||Bx||^2 = x^\mathsf{T} B^\mathsf{T} B x$
- $\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 = x^T A x$ $A = ?$ $= (x_4 - x_3)^2 + (x_3 - x_2)^2 + (x_2 - x_1)^2$
  $= x_4^2 - 2x_3x_4 + x_3^2 + x_3^2 - 2x_2x_3 + x_2^2 + x_2^2 - 2x_2x_1 + x_1^2$
- $\underbrace{||Fx||^2}_{x^T F^T F x} - \underbrace{||Gx||^2}_{x^T G^T G x} = x^T B x$, $B = F^T F - G^T G$

Every quadratic form can be represented as some $||Fx||^2 - ||Gx||^2$

sets defined by quadratic forms:

- $\{ \, x \mid f(x) = a \, \}$ is called a *quadratic surface*
- $\{ \, x \mid f(x) \leq a \, \}$ is called a *quadratic region*



$$x^T A x = \|Fx\|^2 - \|Gx\|^2$$

A real & symmetric

$$A = Q \Lambda Q^T \to x^T A x = x^T Q \Lambda Q^T x$$
$$= x^T Q(\Lambda_+ - \Lambda_-) Q^T x$$

$$\Lambda = \begin{bmatrix} 2 & & & \\ & 3 & & \\ & & -1 & \\ & & & -5 \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & & & \\ & 3 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}}_{\Lambda_+} - \underbrace{\begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & 5 \end{bmatrix}}_{\Lambda_-} \to \underbrace{\Lambda_+^{1/2}}_{\downarrow} = \begin{bmatrix} \pm\sqrt{2} & & & \\ & \pm\sqrt{3} & & \\ & & 0 & \\ & & & 0 \end{bmatrix}$$

$$(\Lambda_+^{1/2})^2 = \Lambda_+$$

$\boxed{\text{Def. } A^{1/2} \text{ is the matrix s.t. } (A^{1/2})^2 = A}$

$$= x^T Q \Lambda_+ Q^T x - x^T Q \Lambda_- Q^T x$$

$$= x^T Q \Lambda_+^{1/2} \underbrace{\boxed{\Lambda_+^{1/2} Q^T}}_{F} x - x^T Q \Lambda_-^{1/2} \underbrace{\boxed{\Lambda_-^{1/2} Q^T}}_{G} x$$

$$= x^T F^T F x - x^T G^T G x$$

$$= \|Fx\|^2 - \|Gx\|^2$$

Inequalities for quadratic forms

suppose $A = A^T$, $A = Q\Lambda Q^T$ with eigenvalues sorted so $\lambda_1 \geq \cdots \geq \lambda_n$ then

$$\boxed{x^T A x \leq \lambda_1 x^T x} = \lambda_{max} \|x\|^2$$

because

$$\begin{aligned}
x^T A x &= x^T Q \Lambda Q^T x \\
&= (Q^T x)^T \Lambda (Q^T x) \\
&= \sum_{i=1}^n \lambda_i (q_i^T x)^2 \\
&\leq \lambda_1 \sum_{i=1}^n (q_i^T x)^2 \\
&= \lambda_1 \|x\|^2
\end{aligned}$$

$\lambda_1 (q_1^T x)^2 + \cdots + \lambda_n (q_n^T x)^2$

$\leq \lambda_1 \quad '' \quad + \cdots + \lambda_1 \, (\, '' \,)^2$

$\|Q^T x\|^2 = x^T \underbrace{Q Q^T}_{I} x$

Inequalities

becomes equality if $x = v_n$

becomes equality if $x = v_1$

- similar argument shows $x^\top Ax \geq \lambda_n ||x||^2$, so we have

$$\lambda_n x^\top x \leq x^\top Ax \leq \lambda_1 x^\top x$$

λ_{min} λ_{max}

$$\to \lambda_n < \frac{x^T A x}{x^T x} < \lambda_1$$

- sometimes λ_1 is called $\lambda_{\max}$, λ_n is called $\lambda_{\min}$
- note also that

$$q_1^\top Aq_1 = \lambda_1 ||q_1||^2, \qquad q_n^\top Aq_n = \lambda_n ||q_n||^2,$$

so the inequalities are tight

Positive semidefinite and positive definite matrices

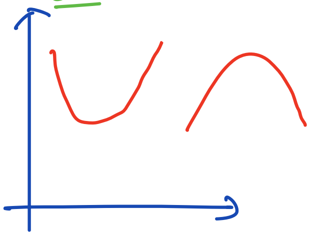
$x = e_i$

$e_i^T A e_i = A_{ii} \geq 0$

suppose $A = A^{\mathsf{T}} \in \mathbb{R}^{n \times n}$

> we say A is **positive semidefinite** if $x^{\mathsf{T}} A x \geq 0$ for all x

$f(x) = ax^2 + bx + c$



- *this is written $A \geq 0$* (and sometimes $A \succeq 0$)
- $A \geq 0$ if and only if $\lambda_{\text{min}}(A) \geq 0$, *i.e.*, all eigenvalues are nonnegative
- **not** the same as $A_{ij} \geq 0$ for all i, j

we say A is *positive definite* if $x^{\mathsf{T}} A x > 0$ for all $x \neq 0$

$f(x) = x^T P x + q^T x + r$

- denoted $A > 0$
- $A > 0$ if and only if $\lambda_{\text{min}}(A) > 0$, *i.e.*, all eigenvalues are positive

Matrix inequalities

- we say A is *negative semidefinite* if $-A \geq 0$ $\iff$ $x^TAx \leq 0$
- we say A is *negative definite* if $-A > 0$ $\iff$ $x^TAx < 0 \quad (x \neq 0)$
- otherwise, we say A is *indefinite*

matrix inequality: if A and B are both symmetric, we use $A < B$ to mean $B - A > 0$.

- many variations, for example $A \geq B$ means $A - B \geq 0$,
- $A > B$ means $x^{\mathsf{T}}Ax > x^{\mathsf{T}}Bx$ for all $x \neq 0$

Matrix inequalities

$A \geq B \iff x^T A x \geq x^T B x$

$C \geq D \iff$ ..

$x^T(A+C)x \geq x^T(B+D)x$

many properties that you'd guess hold actually do, *e.g.*,

- if $A \geq B$ and $C \geq D$, then $A + C \geq B + D$
- if $B \leq 0$ then $A + B \leq A$
- if $A \geq 0$ and $\alpha \geq 0$, then $\alpha A \geq 0$
- $A^2 \geq 0$ $\quad x^T A^2 x = x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$
- if $A > 0$, then $A^{-1} > 0$ $\quad \rightarrow A > 0 \rightarrow \lambda_1, \dots, \lambda_n > 0 \rightarrow 1/\lambda_1, \dots, 1/\lambda_n > 0 \rightarrow A^{-1} > 0$

$\lambda \leftrightarrow A \Rightarrow 1/\lambda \leftrightarrow A^{-1}$: $Av = \lambda v \rightarrow v = \lambda A^{-1} v \rightarrow A^{-1} v = 1/\lambda\, v$

matrix inequality is only a *partial order*: we can have

$$A \not\geq B, \qquad B \not\geq A \qquad\qquad A - B \text{ indefinite}$$

(such matrices are called *incomparable*)