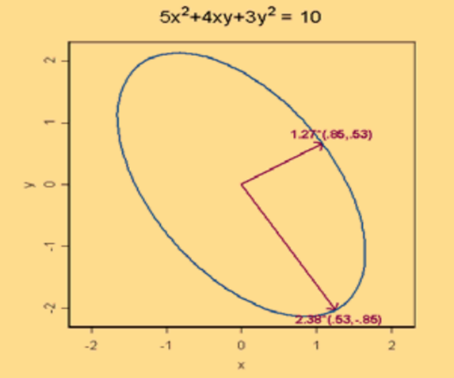
§7.1 Diagonalization of Symmetric Matrices

Equation for an ellipse: $5\,x^2 + 4\,x\,y + 3\,y^2 = 10,$

In matrix form: $\mathbf{x}^T A\,\mathbf{x} = 10, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}.$



Find the major/minor axes and their lengths.

Equation for an ellipse: $\mathbf{x}^T A\,\mathbf{x} = 10, \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}.$

- Eigenvalues of A: $\lambda_1 = 4 - \sqrt{5}, \lambda_2 = 4 + \sqrt{5}$; eigenvectors are

$$A\,\mathbf{v}_1 = \lambda_1\,\mathbf{v}_1, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ \sqrt{5}+1 \end{bmatrix}, \quad A\,\mathbf{v}_2 = \lambda_2\,\mathbf{v}_2, \quad \mathbf{v}_2 = \begin{bmatrix} \sqrt{5}+1 \\ 2 \end{bmatrix}.$$

- $\mathbf{v}_1$ and $\mathbf{v}_2$ are <u>orthogonal</u>: $\mathbf{v}_1^T\,\mathbf{v}_2 = 0$:
- $Q \stackrel{def}{=} \left(\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \right)$ is orthogonal matrix: $Q^{-1} = Q^T$.

$$A = Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^T. \quad \text{Change coordinates: } \begin{bmatrix} u \\ v \end{bmatrix} \stackrel{def}{=} Q^T\mathbf{x}$$

$$10 = \mathbf{x}^T A\,\mathbf{x} = \begin{bmatrix} u \\ v \end{bmatrix}^T \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \left(4 - \sqrt{5}\right)\, u^2 + \left(4 + \sqrt{5}\right)\, v^2.$$

- major/minor axes = Columns of Q,

axis lengths $= \sqrt{\frac{10}{4-\sqrt{5}}}, \sqrt{\frac{10}{4+\sqrt{5}}}$.

Diagonalization of Symmetric Matrices

Let $A \in \mathcal{R}^{n \times n}$ be a symmtric matrix.

Thm 1. Any two real eigenvectors pertaining to two distinct real eigenvalues of A are orthogonal.

Diagonalization of Symmetric Matrices

Let $A \in \mathcal{R}^{n \times n}$ be a symmtric matrix.

Thm 1. Any two real eigenvectors pertaining to two distinct real eigenvalues of A are orthogonal.

PROOF: Let λ_1 and λ_2 be distinct eigenvalues of A, with

$$A\,\mathbf{v}_1 = \lambda_1\,\mathbf{v}_1, \quad A\,\mathbf{v}_2 = \lambda_2\,\mathbf{v}_2.$$

$$\text{so that } \lambda_1\,\mathbf{v}_2^T\,\mathbf{v}_1 = \mathbf{v}_2^T\,(A\,\mathbf{v}_1) = (A\,\mathbf{v}_2)^T\,\mathbf{v}_1 = \lambda_2\,\mathbf{v}_2^T\,\mathbf{v}_1.$$

This implies $(\lambda_2 - \lambda_1)\,\mathbf{v}_2^T\,\mathbf{v}_1 = 0$, or $\mathbf{v}_2^T\,\mathbf{v}_1 = 0$.

Diagonalization of Symmetric Matrices: Main Theorem

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & & \\ & & \end{pmatrix} Q^T.$$

Diagonalization of Symmetric Matrices: Main Theorem

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown \end{pmatrix} Q^T.$$

Note: Assume $A = Q\,D\,Q^T$ with $Q = (\mathbf{q}_1, \cdots, \mathbf{q}_n)$ orthogonal, and $D = \mathbf{diag}\,(d_1, \cdots, d_n)$ diagonal. Then $A\,Q = Q\,D$,

$$A\,(\mathbf{q}_1, \cdots, \mathbf{q}_n) = (\mathbf{q}_1, \cdots, \mathbf{q}_n)\,\mathbf{diag}\,(d_1, \cdots, d_n) = (d_1\,\mathbf{q}_1, \cdots, d_n\,\mathbf{q}_n)\,.$$

- Therefore

$$A\,\mathbf{q}_j = d_j\,\mathbf{q}_j, \quad j = 1, \cdots, n.$$

<u>A has n real eigenvalues with n orthonormal eigenvectors.</u>

Diagonalization of Symmetric Matrices: Main Theorem

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown & \\ & \\ & \end{pmatrix} Q^T.$$

Note: Assume $A = Q\,D\,Q^T$ with $Q = (\mathbf{q}_1, \cdots, \mathbf{q}_n)$ orthogonal, and $D = \mathbf{diag}\,(d_1, \cdots, d_n)$ diagonal. Then $A\,Q = Q\,D$,

$$A\,(\mathbf{q}_1, \cdots, \mathbf{q}_n) = (\mathbf{q}_1, \cdots, \mathbf{q}_n)\,\mathbf{diag}\,(d_1, \cdots, d_n) = (d_1\,\mathbf{q}_1, \cdots, d_n\,\mathbf{q}_n)\,.$$

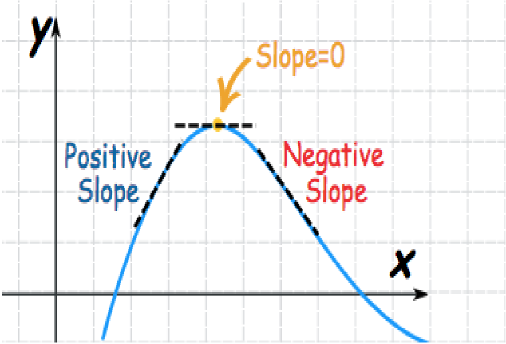
- Therefore

$$A\,\mathbf{q}_j = d_j\,\mathbf{q}_j, \quad j = 1, \cdots, n.$$

<u>A has n real eigenvalues with n orthonormal eigenvectors.</u>

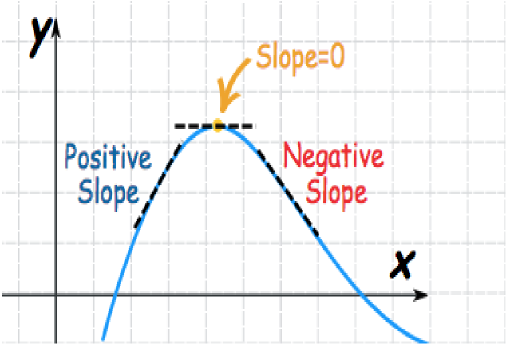
Will prove theorem with Calculus+material from §7.1-7.3 in MIXED order.

Single Variable Calculus: Find MAXIMUM of function (I)



- Before MAXIMUM, function gets bigger with positive slope.
- After MAXIMUM, function gets smaller with negative slope.
- At MAXIMUM, function has 0 slope: $\boxed{f'(x) = 0.}$

Single Variable Calculus: Find MAXIMUM of function (II)



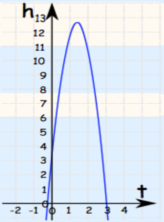
Thm Let $f(x) \in \mathcal{C}[a, b]$ be continuously differentiable, then there exists $x^* \in [a, b]$ so that

$$M = \mathbf{max}_{x \in [a,b]} f(x) = f(x^*).$$

- If $x^* \neq a$ and $x^* \neq b$, then $f'(x^*) = 0$.

Single Variable Calculus: Find MAXIMUM of function (III)

Example: A ball is thrown in the air. Its height at any time t is given by:



$$h = 3 + 14t - 5t^2$$

What is its maximum height?

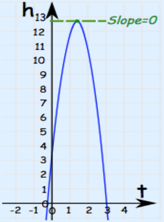
Using derivatives we can find the slope of that function:

$$\frac{d}{dt}h = 0 + 14 - 5(2t)$$
$$= 14 - 10t$$

(See below this example for how we found that derivative.)

Now find when the **slope is zero**:

- $14 - 10t = 0$
- $10t = 14$
- $t = 14 / 10 = \mathbf{1.4}$



The slope is zero at **t = 1.4 seconds**

And the height at that time is:

- $h = 3 + 14 \times 1.4 - 5 \times 1.4^2$
- $h = 3 + 19.6 - 9.8 = \mathbf{12.8}$

And so:

The maximum height is **12.8 m** (at t = 1.4 s)

Muti-Variable Calculus: Gradient

- Muti-Variable function $f(\mathbf{x}) \in \mathcal{R}$ for $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{R}^n$.
- Gradient of $f(\mathbf{x})$ is

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

- At $\textsc{Maximum}$, function has $\mathbf{0}$ slope: $\boxed{\nabla f(\mathbf{x}) = \mathbf{0}.}$

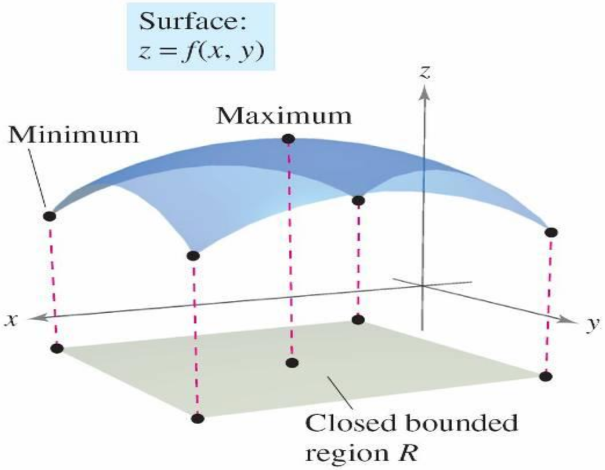
Muti-Variable Calculus: Find MAXIMUM of function (I)

Thm Let function $f(\mathbf{x}) \in \mathcal{R}$ be continuously differentiable for $\mathbf{x}$ in closed region $R \subset \mathcal{R}^n$. Then there exists $\mathbf{x}^* \in R$ so that

$$M = \mathbf{max}_{\mathbf{x} \in R} f(\mathbf{x}) = f(\mathbf{x}^*).$$

- If $\mathbf{x}^*$ is not on boundary of R, then $\nabla f(\mathbf{x}^*) = 0$.

Muti-Variable Calculus: Find MAXIMUM of function (II)



R contains point(s) at which $f(x, y)$ is a minimum and point(s) at which $f(x, y)$ is a maximum.

Muti-Variable Calculus: Find MAXIMUM of function (III)

For a symmetric matrix $A \in \mathcal{R}^{n \times n}$, define closed region

$$R = \{ \mathbf{x} \in \mathcal{R}^n \ \mid \ \|\mathbf{x}\| = 1 \}$$

and continuously differentiable function $\quad f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$.

There must exist $\mathbf{x}^* \in R$ so that

$$M = \mathbf{max}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = (\mathbf{x}^*)^T \, A \, (\mathbf{x}^*) \, .$$

$$\text{Since} \qquad M = \mathbf{max}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = \mathbf{max}_{\|\mathbf{x}\| \neq \mathbf{0}} \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

$$\text{There exists } \mathbf{x}^* \neq \mathbf{0} \quad \text{so} \quad M = \frac{(\mathbf{x}^*)^T \, A \, (\mathbf{x}^*)}{(\mathbf{x}^*)^T \, (\mathbf{x}^*)}.$$

Muti-Variable Calculus: Find MAXIMUM of function (III)

For a symmetric matrix $A \in \mathcal{R}^{n \times n}$, define closed region

$$R = \{\mathbf{x} \in \mathcal{R}^n \ \mid \ \|\mathbf{x}\| = 1\}$$

and continuously differentiable function $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$.

There must exist $\mathbf{x}^* \in R$ so that

$$M = \mathbf{max}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = (\mathbf{x}^*)^T A (\mathbf{x}^*).$$

Since $M = \mathbf{max}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} = \mathbf{max}_{\|\mathbf{x}\| \neq \mathbf{0}} \dfrac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$.

There exists $\mathbf{x}^* \neq \mathbf{0}$ so $M = \dfrac{(\mathbf{x}^*)^T A (\mathbf{x}^*)}{(\mathbf{x}^*)^T (\mathbf{x}^*)}$.

Equation $\nabla \left(\dfrac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right) = \mathbf{0}$ has solution.

Gradient Calculus, with Chain rule (I)

By chain rule,

$$\nabla\left(\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\right) = \frac{\nabla\left(\mathbf{x}^T A\mathbf{x}\right)}{\mathbf{x}^T\mathbf{x}} - \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\,\frac{\nabla\left(\mathbf{x}^T\mathbf{x}\right)}{\mathbf{x}^T\mathbf{x}}.$$

With $\mathbf{x}^T\mathbf{x} = x_1^2 + \cdots + x_n^2$,

$$\nabla\left(\mathbf{x}^T\mathbf{x}\right) = \begin{pmatrix} \frac{\partial\left(x_1^2+\cdots+x_n^2\right)}{\partial x_1} \\ \vdots \\ \frac{\partial\left(x_1^2+\cdots+x_n^2\right)}{\partial x_n} \end{pmatrix} = 2\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 2\,\mathbf{x}.$$

Gradient Calculus, with Chain rule (II)

$$\text{With} \quad \mathbf{x}^T A \mathbf{x} = \begin{array}{l} \quad x_1 \left(a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1n}\, x_n\right) \\ + x_2 \left(a_{21}\, x_1 + a_{22}\, x_2 + \cdots + a_{2n}\, x_n\right) + \cdots \\ + x_n \left(a_{n1}\, x_1 + a_{n2}\, x_2 + \cdots + a_{nn}\, x_n\right) \end{array}$$

$$\begin{aligned} \frac{\partial \left(\mathbf{x}^T A \mathbf{x}\right)}{\partial x_1} &= \left(a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1n}\, x_n\right) + x_1\, a_{11} + x_2\, a_{21} + \cdots + x_n\, a_{n1} \\ &= 2 \left(a_{11}\, x_1 + a_{12}\, x_2 + \cdots + a_{1n}\, x_n\right) = 2 \left(A \mathbf{x}\right)_1 , \\ \frac{\partial \left(\mathbf{x}^T A \mathbf{x}\right)}{\partial x_j} &= 2 \left(a_{j1}\, x_1 + a_{j2}\, x_2 + \cdots + a_{jn}\, x_n\right) = 2 \left(A \mathbf{x}\right)_j , \quad j = 1, \cdots, n. \end{aligned}$$

$$\text{With} \quad \nabla \left(\mathbf{x}^T A \mathbf{x}\right) = \begin{pmatrix} \frac{\partial\left(\mathbf{x}^T A \mathbf{x}\right)}{x_1} \\ \vdots \\ \frac{\partial\left(\mathbf{x}^T A \mathbf{x}\right)}{x_n} \end{pmatrix} = 2 A \mathbf{x}.$$

Muti-Variable Calculus: Find MAXIMUM of function (IV)

By chain rule,

$$\nabla\left(\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\right)=\frac{\nabla\left(\mathbf{x}^T A\mathbf{x}\right)}{\mathbf{x}^T\mathbf{x}}-\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\,\frac{\nabla\left(\mathbf{x}^T\mathbf{x}\right)}{\mathbf{x}^T\mathbf{x}}.$$

But

$$\nabla\left(\mathbf{x}^T\mathbf{x}\right)=2\,\mathbf{x},\quad \nabla\left(\mathbf{x}^T A\mathbf{x}\right)=2\,A\mathbf{x}.$$

Therefore there is solution to

$$\nabla\left(\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\right)=\frac{2\,A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}-\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}\,\frac{2\,\mathbf{x}}{\mathbf{x}^T\mathbf{x}}=\mathbf{0},$$

$$\text{or}\quad A\mathbf{x}=\lambda\,\mathbf{x},\quad \text{for } \mathbf{x}\neq\mathbf{0}.\quad (\text{with eigenvalue } \lambda=\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}.)$$

Therefore $M=\mathbf{max}_{\|\mathbf{x}\|\neq\mathbf{0}}\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$ must be an eigenvalue.

Eigenvector-induced Orthonormal Basis

Let λ be eigenvalue of A with eigenvector $\mathbf{v}$: $A\,\mathbf{v} = \lambda\,\mathbf{v}$.

- We extend $\mathbf{v}$ into a basis for $\mathbb{R}^n$: $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ with $\mathbf{v}_1 = \mathbf{v}$.
- Use Gram-Schmidt to obtain an orthogonal basis for $\mathbb{R}^n$: $\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \cdots, \widehat{\mathbf{v}}_n$ with $\widehat{\mathbf{v}}_1 = \mathbf{v}$.
- Define orthogonal matrix $U \in \mathbb{R}^{n \times n}$

$$U \stackrel{def}{=} \left(\frac{\widehat{\mathbf{v}}_1}{\|\widehat{\mathbf{v}}_1\|}, \cdots, \frac{\widehat{\mathbf{v}}_n}{\|\widehat{\mathbf{v}}_n\|} \right) \stackrel{def}{=} (\mathbf{u}_1, \cdots, \mathbf{u}_n)$$

- λ is eigenvalue of A with UNIT eigenvector $\mathbf{u}_1$: $A\,\mathbf{u}_1 = \lambda\,\mathbf{u}_1$; columns of U orthonormal basis for $\mathbb{R}^n$.

Eigenvector-induced Orthonormal Basis: Example

Matrix $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \in \mathcal{R}^{3\times 3}$ is symmetric with

eigenvalue $\lambda = 6$ and eigenvector $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- An orthogonal basis for $\mathbb{R}^3$:

$$\widehat{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \widehat{\mathbf{v}}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad , \widehat{\mathbf{v}}_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

- Define orthogonal matrix $U \in \mathbb{R}^{n\times n}$

$$U \stackrel{def}{=} \left(\frac{\widehat{\mathbf{v}}_1}{\sqrt{3}}, \frac{\widehat{\mathbf{v}}_2}{\sqrt{2}}, \frac{\widehat{\mathbf{v}}_3}{\sqrt{6}} \right)$$

- λ is eigenvalue of A with UNIT eigenvector

$$\frac{\widehat{\mathbf{v}}_1}{\sqrt{3}} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & & \\ & & \end{pmatrix} Q^T.$$

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown \end{pmatrix} Q^T.$$

Proof:

- By induction on n. Assume theorem true for $n-1$.
- Let λ be eigenvalue of A with UNIT eigenvector $\mathbf{u}$: $A\,\mathbf{u} = \lambda\,\mathbf{u}$.
- We extend $\mathbf{u}$ into an orthonormal basis for $\mathbb{R}^n$: $\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are unit, mutually orthogonal vectors.
- $$U \stackrel{def}{=} (\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n) \stackrel{def}{=} \left(\mathbf{u}, \widehat{U}\right) \in \mathbb{R}^{n \times n} \quad \text{is orthogonal.}$$
  $$U^T A\,U = \begin{pmatrix} \mathbf{u}^T \\ \widehat{U}^T \end{pmatrix} \left(A\,\mathbf{u}, A\,\widehat{U}\right) = \begin{pmatrix} \mathbf{u}^T(A\,\mathbf{u}) & \mathbf{u}^T\left(A\,\widehat{U}\right) \\ \widehat{U}^T(A\,\mathbf{u}) & \widehat{U}^T\left(A\,\widehat{U}\right) \end{pmatrix}$$
  $$= \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{U}^T\left(A\,\widehat{U}\right) \end{pmatrix}.$$
- Matrix $\widehat{U}^T\left(A\,\widehat{U}\right) \in \mathbb{R}^{(n-1)\times(n-1)}$ is symmetric.

Thm: A matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^{n\times n}$ and an orthogonal matrix Q so that $A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & \ddots & \\ & & \ddots \end{pmatrix} Q^T$.

Thm: A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q so

that $A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown \end{pmatrix} Q^T.$

Proof:

- $$U^T A\,U = \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}.$$

- By induction, there exist diagonal matrix $\widehat{D}$ and orthogonal matrix $\widehat{Q} \in \mathbb{R}^{(n-1)\times(n-1)}$,
$$\widehat{U}^T \left(A\,\widehat{U}\right) = \widehat{Q}\,\widehat{D}\,\widehat{Q}^T.$$

- therefore

$$U^T A\,U = \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{Q}\,\widehat{D}\,\widehat{Q}^T \end{pmatrix}.$$

$$A = \left(U \begin{pmatrix} 1 & \\ & \widehat{Q} \end{pmatrix}\right) \begin{pmatrix} \lambda & \\ & \widehat{D} \end{pmatrix} \left(U \begin{pmatrix} 1 & \\ & \widehat{Q} \end{pmatrix}\right)^T \stackrel{def}{=} Q\,D\,Q^T.$$

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric, then

$$M \stackrel{def}{=} \max_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} \quad \text{is the greatest eigenvalue of } A,$$

$$m \stackrel{def}{=} \min_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} \quad \text{is the least eigenvalue of } A.$$

Proof: Write $A = Q\,D\,Q^T$, with orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and diagonal matrix $D = \mathbf{diag}\,(\lambda_1, \cdots, \lambda_n)$ with eigenvalues.

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric, then

$$M \stackrel{def}{=} \mathbf{max}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} \quad \text{is the greatest eigenvalue of } A,$$
$$m \stackrel{def}{=} \mathbf{min}_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x} \quad \text{is the least eigenvalue of } A.$$

Proof: Write $A = Q\,D\,Q^T$, with orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ and diagonal matrix $D = \mathbf{diag}\,(\lambda_1, \cdots, \lambda_n)$ with eigenvalues.

- Define change of variable $\mathbf{y} = Q^T \mathbf{x}$. Then $\|\mathbf{y}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$,

$$\text{and} \quad M = \mathbf{max}_{\|\mathbf{y}\|=1} \mathbf{y}^T D\,\mathbf{y}, \quad m = \mathbf{min}_{\|\mathbf{y}\|=1} \mathbf{y}^T D\,\mathbf{y}.$$

Proof: Write $A = Q\, D\, Q^T$, $D = \mathbf{diag}\,(\lambda_1, \cdots, \lambda_n)$ with eigenvalues,

and $\quad M = \mathbf{max}_{\|\mathbf{y}\|=1}\, \mathbf{y}^T\, D\, \mathbf{y}, \quad m = \mathbf{min}_{\|\mathbf{y}\|=1}\, \mathbf{y}^T\, D\, \mathbf{y},$ for $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$

$$\mathbf{y}^T\, D\, \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^T \mathbf{diag}\,(\lambda_1, \cdots, \lambda_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda_1\, y_1^2 + \cdots + \lambda_n\, y_n^2.$$

Let $\lambda_{\mathbf{max}} = \mathbf{max}\,\{\lambda_1, \cdots, \lambda_n\} = \lambda_{\ell_1}$, $\lambda_{\mathbf{min}} = \mathbf{min}\,\{\lambda_1, \cdots, \lambda_n\} = \lambda_{\ell_2}$, then

$$\lambda_{\mathbf{min}}\, \left(y_1^2 + \cdots + y_n^2\right) \le \lambda_1\, y_1^2 + \cdots + \lambda_n\, y_n^2 \le \lambda_{\mathbf{max}}\, \left(y_1^2 + \cdots + y_n^2\right),$$

$$\text{or,} \quad \lambda_{\mathbf{min}}\, \|\mathbf{y}\|^2 \le \mathbf{y}^T\, D\, \mathbf{y} \le \lambda_{\mathbf{max}}\, \|\mathbf{y}\|^2.$$

$$\text{So for all } \|\mathbf{y}\| = 1, \quad \lambda_{\mathbf{min}} \le m \le \mathbf{y}^T\, D\, \mathbf{y} \le M \le \lambda_{\mathbf{max}}.$$

Proof: For $D = \mathbf{diag}(\lambda_1, \cdots, \lambda_n)$ with eigenvalues,

$$M = \mathbf{max}_{\|\mathbf{y}\|=1} \mathbf{y}^T D \mathbf{y}, \quad m = \mathbf{min}_{\|\mathbf{y}\|=1} \mathbf{y}^T D \mathbf{y}, \text{ for } \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

$$\mathbf{y}^T D \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

Let $\lambda_{\mathbf{max}} = \mathbf{max}\{\lambda_1, \cdots, \lambda_n\} = \lambda_{\ell_1}$, $\lambda_{\mathbf{min}} = \mathbf{min}\{\lambda_1, \cdots, \lambda_n\} = \lambda_{\ell_2}$, then

$$\text{for all } \|\mathbf{y}\| = 1, \quad \lambda_{\mathbf{min}} \le m \le \mathbf{y}^T D \mathbf{y} \le M \le \lambda_{\mathbf{max}}.$$

- Let $\mathbf{e}_j$ be the j^{th} column of the identity.
  - Choose $\mathbf{y} = \mathbf{e}_{\ell_1}$, then $M \ge \mathbf{y}^T D \mathbf{y} = \lambda_{\mathbf{max}}$.
  - Choose $\mathbf{y} = \mathbf{e}_{\ell_2}$, then $m \le \mathbf{y}^T D \mathbf{y} = \lambda_{\mathbf{min}}$.
- Therefore $M = \lambda_{\mathbf{max}}$, $m = \lambda_{\mathbf{min}}$.

Thm: Let matrix $A \in \mathbb{R}^{n\times n}$ be symmetric, then

$$M \overset{def}{=} \max_{\|\mathbf{x}\|=1} \mathbf{x}^T A\,\mathbf{x} \quad \text{is the greatest eigenvalue of } A,$$

$$m \overset{def}{=} \min_{\|\mathbf{x}\|=1} \mathbf{x}^T A\,\mathbf{x} \quad \text{is the least eigenvalue of } A.$$

EXAMPLE: Matrix $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \in \mathcal{R}^{3\times 3}$ is symmetric with eigenvalues $\lambda_1 = 6,\ \lambda_2 = 3,\ \lambda_3 = 1$ and unit eigenvectors

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},\ \mathbf{u}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix},\ \mathbf{u}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Therefore

$$M = \mathbf{u}_1^T A\,\mathbf{u}_1 = 6, \quad m = \mathbf{u}_3^T A\,\mathbf{u}_3 = 1.$$