LINEAR ALGEBRA

CHAPTER 6. EIGENVALUES AND EIGENVECTORS

Prof. Cheolsoo Park

BCML Bio Computing & Machine Learning Lab

INTRODUCTIONS TO EIGENVALUES

- A matrix is square
- Function *f(x) :* $A\mathrm{x}$

  , where a vector x, parallel to $A\mathrm{x}$, is called an **eigenvector**

$$A\mathrm{x} = \lambda\mathrm{x}$$

$$\lambda \text{ is an eigenvalue}$$

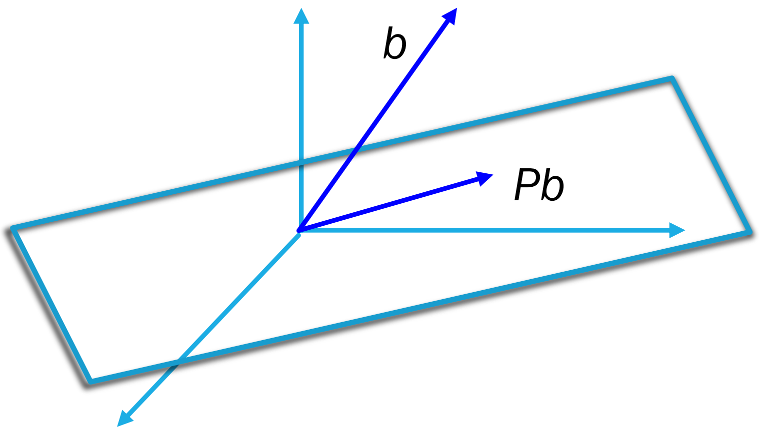
- What if λ is zero, then $A\mathrm{x} = 0$ ← nullspace

  ← If A is singular, then $\lambda = 0$

- Apart from $\lambda = 0$, we need to know all eigenvalues λ

INTRODUCTIONS TO EIGENVALUES

- Example : Projection matrix P



What are x's and λ's for a projection matrix?

Any x in a plane : Px = x

➔ Eigenvector : x

Eigenvalue : $\boldsymbol{\lambda = 1}$

Any x' ⊥ plane : Px' = 0x', $\boldsymbol{\lambda = 0}$

- If the eigenvalue λ is zero, then $Ax = 0x$.

  And if x is a non-zero eigenvector, then *A* should be singular)

- If A is the identity matrix, every vector has $Ax = x$

  → all vectors are eigenvectors of I and all eigenvalues are $\lambda = 1$

INTRODUCTIONS TO EIGENVALUES

Example) Permutation matrix

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

$$\text{for } \mathrm{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathrm{Ax}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathrm{x}_1 \text{ and } \lambda = 1$$

$$\text{for } \mathrm{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathrm{Ax}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\mathrm{x}_2 \text{ and } \lambda = -1$$

THE EQUATION FOR THE EIGENVALUES

- 'Trace' of a square matrix : sum of diagonal
- Sum of eigenvalues equals the sum of the diagonals, the trace
- Eigenvectors are ~~perpendicular~~ : $x_1 \cdot x_2 = 0$

- How to solve $A\mathrm{x} = \lambda\mathrm{x}$

Rewrite: $(A - \lambda I)\mathrm{x} = 0$

In $(A - \lambda I)$, the diagonal terms are shifted by λ

And $(A - \lambda I)$ should be **<u>singular</u>** ➔ $\det(A - \lambda I) = 0$ → find λ first

THE EQUATION FOR THE EIGENVALUES

Ex) $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 1 = 0 \rightarrow \lambda$=4 and 2

4+2=trace 3+3

1. find the nullspace of $A - 4\lambda = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$

➔ $\mathrm{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

* RN=0 ➔ N=$\begin{bmatrix} -F \\ I \end{bmatrix}$

2. find the nullspace of $A - 2\lambda = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

➔ $\mathrm{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\mathrm{x}_1 \cdot \mathrm{x}_2 = 0$

THE EQUATION FOR THE EIGENVALUES

- $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ → ($\mathrm{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, λ_1=4) and ($\mathrm{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, λ_2=2)

  $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ → ($\mathrm{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, λ_1=1) and ($\mathrm{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, λ_2=-1)

← Same eigenvectors, and increases of eigenvalues by 3, which is the same as the trace's increase

➔ If $A\mathrm{x}=\lambda\mathrm{x}$, then $(A + 3I)\mathrm{x} = \lambda\mathrm{x} + 3\mathrm{x} = (\lambda + 3)\mathrm{x}$

THE EQUATION FOR THE EIGENVALUES

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

A is singular : for $A\mathbf{x} = 0$, A's nullspace exists → $\mathrm{A}\mathbf{x} = \lambda\mathbf{x}$, and $\lambda_1 = 0$,

Sum of eigenvalues equals the sum of diagonals→$\lambda_2 = -3$

This can also be calculated using $|A - \lambda\, I| = 0$

Eigenvector for $\lambda_1 = 0$: $A\mathrm{x}_1 = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} X_1 = 0X_1 \rightarrow X_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Eigenvector for $\lambda_2 = -3$: $(A - (-3)I)\mathrm{x}_2 = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} X_2 = 0 \rightarrow X_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

PROPERTY OF EIGENVALUES

- If $A\mathrm{x} = \lambda\mathrm{x}$ and $B\mathrm{x} = \alpha\mathrm{x}$, then $(A + B)\mathrm{x} = (\lambda + \alpha)\mathrm{x}$

  But it's wrong!

  It is because their eigenvectors might not be the same

  So $A\mathrm{x} = \lambda\mathrm{x}$ and $B\mathrm{y} = \alpha\mathrm{y}$
- In addition, $(AB)\mathrm{x} = (\lambda\alpha)\mathrm{x}$ is also wrong!
- No linearity in eigenvalues!

IMAGINARY EIGENVALUES

Example) 90° rotation matrix $Q = \begin{bmatrix} \cos(90°) & -sin(90°) \\ sin(90°) & \cos(90°) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

sum of the eigenvalues is zero because of the trace

determinant of Q = 1 = $\lambda_1\lambda_2$ ←this is the property of eigenvalue

$\det(Q - \lambda I) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \rightarrow \lambda_1 = i \text{ and } \lambda_2 = -i$

- ***Symmetric matrix has real number eigenvalues***

  Anti-symmetric matrix has imaginary number eigenvalues

Example) Triangular matrix A= $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$

$$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -1 \\ 0 & 3-\lambda \end{vmatrix} = (3-\lambda)(3-\lambda)$$

$$\lambda_1 = 3, \lambda_2 = 3$$

$$(A - \lambda I)\mathrm{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\mathrm{x} = 0$$

← $\mathrm{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and no 2nd independent eigenvector

DIAGONALIZING A MATRIX

Diagonalization Suppose the n by n matrix A has n linearly independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Put them into the columns of an ***eigenvector matrix*** S. Then $S^{-1}AS$ is the ***eigenvalue matrix*** Λ:

$Ax = \lambda x$

Eigenvector matrix S
Eigenvalue matrix Λ

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \qquad (1)$$

Proof Multiply A times its eigenvectors, which are the columns of S. The first column of AS is $A\boldsymbol{x}_1$. That is $\lambda_1 \boldsymbol{x}_1$. Each column of S is multiplied by its eigenvalue λ_i:

A times S
$$AS = A\begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & \cdots & \lambda_n\boldsymbol{x}_n \end{bmatrix}.$$

The trick is to split this matrix AS into S times Λ:

S times Λ
$$\begin{bmatrix} \lambda_1\boldsymbol{x}_1 & \cdots & \lambda_n\boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = S\Lambda.$$

*Key for PCA!!! $Q^T A Q = \Lambda$ and $A = Q\Lambda Q^T$

DIAGONALIZING A MATRIX

- If $Ax = \lambda x$,

  $A^2x = \lambda Ax = \lambda^2 x$

- $A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}$

$$A^k = S\Lambda^k S^{-1}$$

- Theorem

$$A^k \to 0 \; as \; k \to \infty, if \; all \; |\lambda_i| < 1$$

- Eigenvector approach needs n independent eigenvectors
- If we don't have n independent eigenvectors, we can't diagonalize the matrix : S^{-1} should exist→ $S^{-1}AS = \Lambda$

DIAGONALIZING A MATRIX

- A is sure to have n independent eigenvectors and be diagonalizable if all the λ's are different (no repeated λ's)
- Repeated eigenvalues

Ex) $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ → $\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{vmatrix}$ → λ=2 and 2

nullspace of $A - 2I$=$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ → $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

➔ S only have one vector, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which cannot be invertible

→ no S^{-1} and so we cannot have $S^{-1}AS$ for diagonalization

SYMMETRIC MATRICES

- Symmetric matrix, $A = A^T$
  - The eigenvalues are REAL
  - The eigenvectors are **PERPENDICULAR**
- Usual case : $A = S\Lambda S^{-1}$

  Symmetric case : $A = Q\Lambda \mathrm{Q}^{-1}$ (Q is orthonormal→ $Q^T = Q^{-1}$)

$$A = Q\Lambda \mathrm{Q}^{\mathrm{T}}$$

SYMMETRIC MATRICES

- Why the eigenvalues of symmetric matrix A is real?

$$A\mathrm{x} = \lambda\mathrm{x} \Rightarrow A\bar{x} = \bar{\lambda}\bar{\mathrm{x}} \quad (conjugate)$$

$$\Rightarrow A\mathrm{x} = \lambda\mathrm{x} \Rightarrow \bar{\mathrm{x}}^T A\,\mathrm{x} = \lambda\bar{\mathrm{x}}^T\mathrm{x}$$

$$\Rightarrow \bar{\mathrm{x}}^T A^T = \bar{\mathrm{x}}^T A = \bar{\mathrm{x}}^T\bar{\lambda} \Rightarrow \bar{\mathrm{x}}^T A\,\mathrm{x} = \bar{\mathrm{x}}^T\bar{\lambda}\,\mathrm{x}$$

$$\Rightarrow \lambda\underline{\bar{\mathrm{x}}^T\mathrm{x}} = \bar{\lambda}\,\bar{\mathrm{x}}^T\,\mathrm{x} \Rightarrow \lambda = \bar{\lambda};\ \boldsymbol{\lambda}\ \textbf{is real!}$$

$$\bar{\mathrm{x}}^T\,\mathrm{x} = [\bar{\mathrm{x}}_1\ \bar{\mathrm{x}}_2 \ldots \bar{\mathrm{x}}_n]\begin{bmatrix}\mathrm{x}_1\\ \mathrm{x}_2\\ \vdots\\ \mathrm{x}_n\end{bmatrix} = \bar{\mathrm{x}}_1\mathrm{x}_1 + \bar{\mathrm{x}}_2\mathrm{x}_2 + \cdots$$

$\leftarrow$ $positive\ real\ number, e.g.\ (a+ib)(a-ib)$

SYMMETRIC MATRICES

- **Orthogonal Eigenvectors**
  - Eigenvectors of a real symmetric matrix (when they correspond to different λ's) are **always perpendicular**

Proof Suppose $A\boldsymbol{x} = \lambda_1\boldsymbol{x}$ and $A\boldsymbol{y} = \lambda_2\boldsymbol{y}$. We are assuming here that $\lambda_1 \neq \lambda_2$. Take dot products of the first equation with $\boldsymbol{y}$ and the second with $\boldsymbol{x}$:

$$\text{Use } A^{\mathrm{T}} = A \qquad (\lambda_1\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}\lambda_2\boldsymbol{y}. \qquad (3)$$

The left side is $\underline{\boldsymbol{x}^{\mathrm{T}}\lambda_1\boldsymbol{y}}$, the right side is $\underline{\boldsymbol{x}^{\mathrm{T}}\lambda_2\boldsymbol{y}}$. Since $\lambda_1 \neq \lambda_2$, this proves that $\underline{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = 0}$. The eigenvector $\boldsymbol{x}$ (for λ_1) is perpendicular to the eigenvector $\boldsymbol{y}$ (for λ_2).

SYMMETRIC MATRICES

- 'Symmetric' is key property for a good matrix, meaning real eigenvalues and perpendicular eigenvectors

$$A = A^T \Rightarrow A = Q\Lambda Q^T$$

- For a complex matrix A

$$A = \bar{A}^T \text{ (Hermitian matrix)}$$

the Hermitian matrix can also have real eigenvalues and perpendicular eigenvectors

SYMMETRIC MATRICES

- ***Principal component analysis***
  - Covariance matrix $C = \mathrm{x}^{\mathrm{T}}\mathrm{x}$, which is real and symmetric

$$C = Q\Lambda Q^T \;\Rightarrow\; Q^T C Q = \Lambda$$

  - The covariance matrix has been diagonalized

    ➔ no correlation among different $x's$

  - $Q^T C Q \Rightarrow \; Q^T \mathrm{x}^{\mathrm{T}}\mathrm{x} Q$: $\mathrm{x}Q$ is a projected matrix of x into Q
  - This can be a spatial filter for a multichannel data x

SYMMETRIC MATRICES

- $A = A^T \Rightarrow A = Q\Lambda Q^T$

$$= [q_1\ q_2\ \dots] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \ddots \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \end{bmatrix}$$

$$= \lambda_1 \underline{q_1 q_1^T} + \lambda_2 q_2 q_2^T + \cdots$$

Projection matrix, which was $P = \frac{aa^T}{a^Ta}$

The denominator is a scalar since a is a column vector

- Every symmetric matrix is a combination of perpendicular projection matrices

BCML Bio Computing & Machine Learning Lab

SYMMETRIC MATRICES

- Signs of pivots of symmetric matrices are the same as the signs of eigenvalues

  ➔ # of positive pivots == # of positive eigenvalues

- Product of pivots of symmetric matrices is equal to the product of eigenvalues

  ← both equal to the determinant

POSITIVE DEFINITE MATRIX

- Positive definite matrix : for a symmetric matrix, all eigenvalues are positive
- If all eigenvalues are positive, all the pivots are positive

  (# of positive pivots = # of positive eigenvalues)

Ex) $\begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}$, its first pivot : 5

product of pivots = determinant

➔ the second pivot is $5x = 11 \Rightarrow \frac{11}{5}$

positive pivots 5 and $\frac{11}{5}$ → all eigenvalues are also positive

Double check : $\begin{vmatrix} 5-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix}$ → $\lambda^2 - 8\lambda + 11 = 0 \rightarrow \lambda = 4 \pm \sqrt{5}$

POSITIVE DEFINITE MATRIX

- There are faster way to check all eigenvalues are positive rather than calculating eigenvalues and checking whether they are positive

- Positive definite test for a symmetric matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

1. $\lambda_1 > 0 \; and \; \lambda_2 > 0$

2. $1 \times 1 \; determinant, \begin{bmatrix} |a| & b \\ b & c \end{bmatrix}, = a > 0,$

   $and \; 2 \times 2 \; determinant, \begin{vmatrix} a & b \\ b & c \end{vmatrix}, = ac - b^2 > 0$

3. pivots $a > 0 \; and \; \frac{ac-b^2}{a} > 0$ ← Product of pivots = determinant

4. $\mathrm{x^T A x} > 0$ ← $\mathrm{x^T A x}$ could be a λ_k (x can be one of eigenvector)

POSITIVE DEFINITE MATRIX

Example) $\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$

→ **positive semi definite :**

; it is singular deriving $\lambda_1 = 0$ since the product of eigenvalues equals determinant (==0), and $\lambda_2 = 20$ because the sum of eigenvalues equals trace

; Pivots : 2, only one pivot due to the singular matrix

; $\mathrm{x^T A x} = [x_1\ x_2]\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1\ x_2]\begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 18x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 18x_2^2$

is this $2x_1^2 + 12x_1x_2 + 18x_2^2 > 0$???

$$2(x_1^2 + 6x_1x_2 + 9x_2^2) = 2(x_1 + 3x_2)^2 \geq 0$$

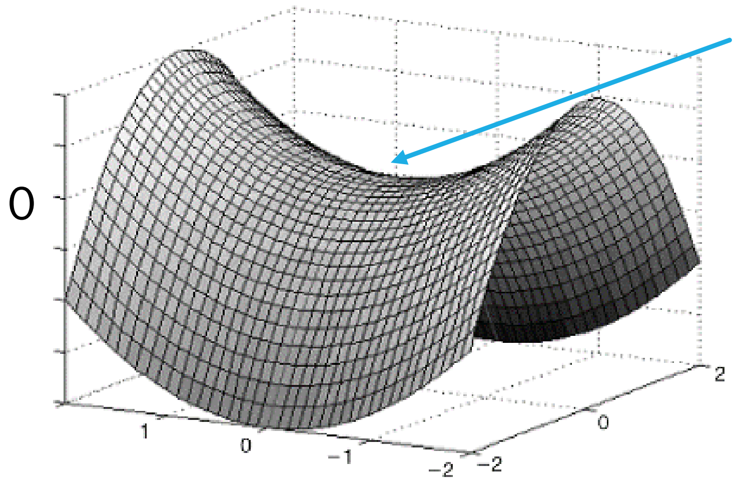
POSITIVE DEFINITE MATRIX

Example) $\begin{bmatrix} 2 & 6 \\ 6 & 7 \end{bmatrix}$ → not a positive definite matrix due to det <0

; $x^T A x = [x_1\ x_2] \begin{bmatrix} 2 & 6 \\ 6 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1\ x_2] \begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 7x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 7x_2^2$

this could be negative

The graph of $f(x, y) = 2x_1^2 + 12x_1x_2 + 7x_2^2$



It has a saddle point!

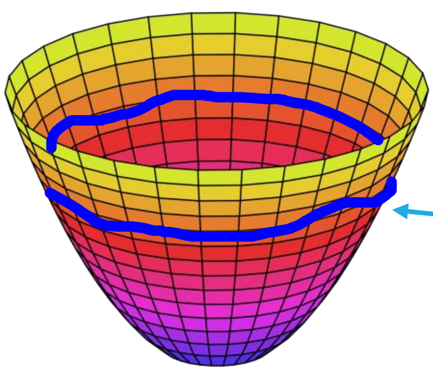
POSITIVE DEFINITE MATRIX

Example) $\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$ → positive definite matrix

because $\lambda_1 \ \lambda_2 > 0$ since det > 0, and $\lambda_1 + \lambda_2 = trace = 22$

; $\mathrm{x^T A x} = [x_1 \ x_2] \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \ x_2] \begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 20x_2 \end{bmatrix} = 2x_1^2 + 12x_1x_2 + 20x_2^2$

since the matrix is positive definite, $\mathrm{x^T A x} > 0$ except at x=0



Or derive

$$2x_1^2 + 12x_1x_2 + 20x_2^2 = 2(x_1 + 3x_2)^2 + 2x_2^2 > 0$$

1= $2x_1^2 + 12x_1x_2 + 20x_2^2 \Rightarrow Ellipse$ *(size increases as λ increases)*

For 3 × 3 matrix, it becomes sphere

1st derivative =0

2nd derivative>0

$*2^{\mathrm{nd}}$ deriv., $\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$, should be positive definite

This is like a positive meaning for a matrix

POSITIVE DEFINITE MATRIX

- LU factorization of A $= \begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$=UL,

  where $U = \begin{bmatrix} 2 & 6 \\ 0 & 2 \end{bmatrix}$ $and\ L = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$

  multiplier

  pivots

  $2(x_1 + 3x_2)^2 + 2x_2^2$ >0

  ➔ positive pivots give sum of squares

  ➔ all the values are positive

  ➔ graph goes up

  ➔ it can be extended to $n \times n$ as well