Plan:

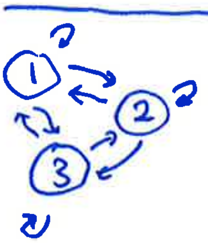
Review: Lecture 5

1. Definiteness of symmetric matrices
2. Unconstrained optimization in several variables
3. Convex and concave functions

{ more on convex/concave fn. in Lecture 7

Review:

Markov chains



$\rightsquigarrow$ $A = (a_{ij})$ transition matrix

$\underline{x}_{t+1} = A\,\underline{x}_t$ state vectors $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

a_{ij}: (j) $\to$ (i) probability

Regular: A^n consists of non-zero entries for some n

(j) $\to \to \cdots \to$ (i) n steps

$$\boxed{\begin{array}{c} a_{ij} > 0 \text{ for all } i,j \\ \Downarrow \\ A \text{ regular} \end{array}}$$

Theory: If A is regular, then there is a unique eigenvector $\underline{v}$ in E_1 ($\lambda = 1$) which is a state vector ($v_1, v_2, v_3 \geq 0$, $v_1 + v_2 + v_3 = 1$), and $\underline{v}$ is the equilibrium state

$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ (for any $\underline{v}_0$, $A^n \cdot \underline{v}_0 \to \underline{v}$ as $n \to \infty$)

Note: If A is not regular, then we have to compute $A^n \cdot \underline{v}_0$ and its limit as $n \to \infty$; it may depend on $\underline{v}_0$.

Definiteness of quadratic forms / symmetric matrices

$f(x_1, \ldots, x_n) = c_{11}x_1^2 + c_{12}x_1x_2 + \ldots \longleftrightarrow$ A symmetric $n\times n$-matrix $\qquad f(\underline{x}) = \underline{x}^T A \underline{x}$

quadratic form

A symm. $\Rightarrow$ A diagonalizable

$\Rightarrow$ n eigenvalues $\underline{\lambda_1, \ldots, \lambda_n}$

Results

- A pos. definite $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n > 0$
- A pos. semi-defn. $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n \geq 0$
- A neg. defn. $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n < 0$
- A neg. semi-defn. $\Leftrightarrow$ $\lambda_1, \ldots, \lambda_n \leq 0$
- A indefinite $\Leftrightarrow$ all other cases (both pos. and neg. eigenvalues)

Fact:

Any quadratic form can be written as

$$f(x_1, \ldots, x_n) = \lambda_1 \cdot u_1^2 + \lambda_2 u_2^2 + \ldots + \lambda_n u_n^2$$

where $u_1, u_2, \ldots, u_n$ are linear expressions in $x_1, \ldots, x_n$.

Ex:

$$4x_1x_2 = (x_1+x_2)^2 - (x_1-x_2)^2$$

$$A = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$$

$$= \underbrace{2}_{\lambda_1} \cdot \Big(\underbrace{\frac{x_1+x_2}{\sqrt{2}}}_{u_1}\Big)^2 \underbrace{-2}_{\lambda_2} \Big(\underbrace{\frac{x_1-x_2}{\sqrt{2}}}_{u_2}\Big)^2$$

$\lambda^2 - 4 = 0$

$\underline{\lambda_1 = 2}, \underline{\lambda_2 = -2}$

indefinite

Leading principal minors:

A $n \times n$ symmetric matrix.

$D_1 = M_{1,1}$

$D_2 = M_{12,12}$

$D_3 = M_{123,123}$

$\vdots$

$D_n = |A|$

D_i = leading principal minor of order i, obtained by keeping rows $1,2,\dots,i$ and cols $1,2,\dots,i$.

Result:

A is pos. defn. $\iff D_1, D_2, \dots, D_n > 0$

A is neg. defn. $\iff D_1 < 0,\ D_2 > 0, \dots$

($D_i < 0$ if i is odd, $D_i > 0$ if i is even)

Ex: $A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$

$f = -x^2 - 2y^2$

$\lambda_1 = -1 \quad \lambda_2 = -2$

$D_1 = \underset{\lambda_1}{-1} < 0$

$D_2 = \underset{\lambda_1 \lambda_2}{2} > 0$

A is neg. defn.

$f(0,0) = 0$ is global max

Ex: $f = x_1^2 + x_2^2 - x_1x_2$

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

$D_1 = 1 > 0$

$D_2 = 1 - (1/4) = 3/4 > 0$

f is pos. defn.

$(0,0)$ global min

Ex: $f = xw - yz$

$A = \begin{pmatrix} 0 & 0 & 0 & 1/2 \\ 0 & 0 & -1/2 & 0 \\ 0 & -1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \end{pmatrix}$

$D_1 = 0$

$D_2 = 0$

$D_3 = 0$

$D_4 = -\frac{1}{2}\left(-\frac{1}{2} \cdot (0 + 1/4)\right)$

$= \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16} > 0$

f is pos. semidefn?

All principal minors:

Δ_i : principal minor of order i, obtained by keeping i rows $(i_1 i_2 \ldots i_i)$ and the same i cols.

$$\Delta_1 = \underset{\substack{\| \\ D_1}}{M_{1,1}},\ M_{2,2},\ M_{3,3},\ M_{4,4} \qquad \leftarrow n=4$$

$$\Delta_2 = \underset{\substack{\| \\ D_2}}{M_{12,12}},\ M_{13,13},\ M_{14,14},$$

$$M_{23,23},\ M_{24,24},\ M_{34,34}$$

$$\Delta_3 = \underset{\substack{\| \\ D_3}}{M_{123,123}},\ M_{124,124},\ M_{134,134},\ M_{234,234}$$

$$\Delta_4 = \underset{\substack{\| \\ D_4}}{|A|}$$

Ex: $A = \begin{pmatrix} 0 & 0 & 0 & 1/2 \\ 0 & 0 & -1/2 & 0 \\ 0 & -1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \end{pmatrix}$

$\Delta_1 = 0,\ 0,\ 0,\ 0$

$\Delta_2 = 0,\ 0,\ -1/4,\ \ldots$ $\leftarrow$ A indefinite

$\Delta_3 = 0,\ \ldots$

$\Delta_4 = 1/16$

$\uparrow$ D_i

Result:

A pos. semidefn. $\iff$ $\Delta_1, \Delta_2, \Delta_3, \Delta_4 \ldots \geq 0$ for all principal minors

A neg. semidefn $\iff$ $\begin{cases} \Delta_i \leq 0 & \text{for all principal minors with } i \text{ odd} \\ \text{and} & \\ \Delta_i \geq 0 & \text{for all principal minors with } i \text{ even} \end{cases}$

A indefinite $\iff$ all other cases

Extra example:

$$f(x,y,z,w) = x^2 + y^2 + 6yz - 4yw + 9z^2 - 12zw + 4w^2$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 3 & 9 & -6 \\ 0 & -2 & -6 & 4 \end{pmatrix}$$

$D_1 = \underline{1}$

$D_2 = \underline{1}$

$D_3 = 1 \cdot \begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} = 1 \cdot (9-9) = \underline{0}$

$D_4 = 1 \cdot \begin{vmatrix} 1 & 3 & -2 \\ 3 & 9 & -6 \\ -2 & -6 & 4 \end{vmatrix} = 1 \cdot (1 \cdot (36-36) - 3 \cdot (12-12) + (-2) \cdot (-18+18))$

$= \underline{0}$

pos. semidef. ?

$\Delta_1 = \underset{\substack{\parallel \\ D_1 = M_{1,1}}}{1}, \underset{\substack{\parallel \\ M_{2,2}}}{1}, \underset{\substack{\parallel \\ M_{3,3}}}{9}, \underset{\substack{\parallel \\ M_{4,4}}}{4} \geq 0 \quad \checkmark$ ok.

$\Delta_2 = \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1}_{D_2 = M_{12,12}}, \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 9 \end{vmatrix} = 9}_{M_{13,13}}, \underbrace{\begin{vmatrix} 1 & 0 \\ 0 & 4 \end{vmatrix} = 4}_{M_{14,14}}, \underbrace{\begin{vmatrix} 1 & 3 \\ 3 & 9 \end{vmatrix} = 0}_{M_{23,23}}, \underbrace{\begin{vmatrix} 1 & -2 \\ -2 & 4 \end{vmatrix} = 0}_{M_{24,24}},$

$\underbrace{\begin{vmatrix} 9 & -6 \\ -6 & 4 \end{vmatrix} = 0}_{M_{34,34}} \quad \geq 0 \quad \checkmark$ ok

$\Delta_3 = \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 9 \end{vmatrix} = 0}_{D_3 = M_{123,123}}, \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{vmatrix} = 0}_{M_{124,124}}, \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 9 & -6 \\ 0 & -6 & 4 \end{vmatrix} = 0}_{M_{134,134}}, \underbrace{\begin{vmatrix} 1 & 3 & -2 \\ 3 & 9 & -6 \\ -2 & -6 & 4 \end{vmatrix} = 0}_{M_{234,234}} \geq 0$ ok ✓

$\Delta_4 = \underbrace{\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 3 & 9 & -6 \\ 0 & -2 & -6 & 4 \end{vmatrix} = 0}_{D_4 = |A|} \geq 0$ ok ✓

Conclusion:

$\Delta_i \geq 0$ for all i

f is pos. semidef.

Note: If $\Delta_i < 0$ for i even, then A is indefinite

If Δ_i's have opposite signs for i odd, then A is indefinite.

Ex: $f = x^2 + y^2 + z^2 + w^2 - 2xw - 2yz$

$f(x,y,z,w)$

$$A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$D_1 = 1 > 0$
$D_2 = 1 > 0$
$D_3 = 1 \cdot 0 = 0$
$D_4 = 1 \cdot (-1) \cdot 0 + 1 \cdot 0 = 0$

f pos. semidefn.?

$\Delta_1 = 1, 1, 1, 1 \geq 0$
$\Delta_2 = 1, 1, 0, 0, 1, 1 \geq 0$
$\Delta_3 = 0, 0, 0, 0 \geq 0$
$\Delta_4 = 0$

f pos. semidefn!

① Reduced rank criterion

Thm: (2016)

If A is an $n \times n$ symmetric matrix of $rk(A) = r < n$, then:

$D_1, D_2, \ldots, D_r > 0 \implies A$ pos. semidefn.

$D_1 < 0, D_2 > 0, \ldots, (-1)^r D_r > 0 \implies A$ neg. semidefn.

for $i = 1, 2, \ldots, r$, we have
$D_i < 0$ for i odd
$D_i > 0$ for i even

Ex:
$$A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$rk(A) = 2$ $\quad r = 2$, $n = 4$

RRC:

$D_1 = 1 > 0$, $D_2 = 1 > 0$ $\Big\}$ A pos. semidefn.

$f(x,y,z,w) = x^2 + y^2 + z^2 + w^2 - 2xw - 2yz$

2 Unconstrained optimization

max/min $\underbrace{f(x_1, x_2, \ldots, x_n)}_{\text{objective fn.}}$ $\leftarrow$ max = global max, min = global min

no constraints

Method:

1 Find stationary pts of f:

A pt. $\underline{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$ is a stationary pt if

FOC: $\to$ $\boxed{f'_{x_1}(\underline{x}^*) = 0, \; f'_{x_2}(\underline{x}^*) = 0, \; \ldots, \; f'_{x_n}(\underline{x}^*) = 0}$

Ex: $f(x,y,z) = x^3 - 3xy + y^3 + z^2$

FOC:
$f'_x = 3x^2 - 3y = 0$ $\Rightarrow \frac{3y}{3} = \frac{3x^2}{3} \Rightarrow y = x^2$

$f'_y = -3x + 3y^2 = 0$ $\quad -3x + 3\cdot(x^2)^2 = 0$

$f'_z = 2z = 0$ $\quad \Rightarrow z = 0$

$-3x + 3x^4 = 0$

$3x(-1 + x^3) = 0$

$x = 0$ or $x^3 = 1$ $\quad x = 1$

$x = \sqrt[3]{1} = 1$

$x = 0, y = 0, z = 0$ or $x = 1, y = 1, z = 0$

Stationary pts: $(x,y,z) = (0,0,0), (1,1,0)$

Fact: If f has partial derivatives, then:

$$\boxed{\mathbf{x}^* \text{ max/min} \implies \mathbf{x}^* \text{ stationary pt.}}$$

$\Downarrow$

Candidates for max/min: stationary pts

max = global max
min = global min

Ex: $f = x^3 - 3xy + y^3 + z^2$

$$\boxed{\begin{aligned}(x,y,z) = {} & (0,0,0), \\ & (1,1,0)\end{aligned}}$$

$f(0,0,0) = 0$

$f(1,1,0) = -1$

② Classify stationary pts

- local max $\left\{ f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for all } \mathbf{x} \text{ close to } \mathbf{x}^* \right.$
- local min $\left\{ f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ for all } \mathbf{x} \text{ close to } \mathbf{x}^* \right.$
- saddle pt. $\left\{ \text{all other cases} \right.$

$$H(f) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots \\ f''_{x_2x_1} & & \ddots \\ \vdots & & \end{pmatrix}$$

Hessian matrix of all second order partial derivatives

Fact: When f is "nice", $H(f)$ is symmetric.

Ex: $f = x^3 - 3xy + y^3 + z^2$

$f'_x = 3x^2 - 3y$

$f'_y = -3x + 3y^2$

$f'_z = 2z$

$f''_{xx} = 6x \quad f''_{xy} = -3 \quad f''_{xz} = 0$

$f''_{yx} = -3 \quad f''_{yy} = 6y \quad f''_{yz} = 0$

$f''_{zx} = 0 \quad f''_{zy} = 0 \quad f''_{zz} = 2$

$$H(f) = \begin{pmatrix} 6x & -3 & 0 \\ -3 & 6y & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$\parallel$

$H(f)(x,y,z)$

Second derivative test:

If $\underline{x}^*$ is a stationary pt. for f, then:

$H(f)(\underline{x}^*)$ positive defn. $\Rightarrow$ $\underline{x}^*$ local min

$H(f)(\underline{x}^*)$ negative defn. $\Rightarrow$ $\underline{x}^*$ local max

$H(f)(\underline{x}^*)$ indefinite $\Rightarrow$ $\underline{x}^*$ saddle point

In all other cases, then the test is inconclusive.

Ex: $f = x^3 - 3xy + y^3 + z^2$

Candidate pts: (stationary pts) $(x,y,z) = (0,0,0), (1,1,0)$

$$H(f) = \begin{pmatrix} 6x & -3 & 0 \\ -3 & 6y & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

$(0,0,0)$: $H(f)(0,0,0) = \begin{pmatrix} 0 & -3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

$D_1 = 0$
$D_2 = -9 < 0$
indefinite
$\Downarrow$
$(0,0,0)$ saddle pt.

$(1,1,0)$: $H(f)(1,1,0) = \begin{pmatrix} 6 & -3 & 0 \\ -3 & 6 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

$D_1 = 6 > 0$
$D_2 = 27 > 0$
$D_3 = 2 \cdot 27 = 54 > 0$
pos. defn
$\Downarrow$
$(1,1,0)$ local min

③ Conclude regarding global max/min

Ex: $f = x^3 - 3xy + y^3 + z^2$ $\quad f(-2,0,0) = -8$

$(0,0,0)$ saddle pt. $\quad f = 0$ $\quad$ no global min

$(1,1,0)$ local min $\quad f = -1$ $\quad$ no global max

Note: ① global max ⇒ local max
global min ⇒ local min

local classification gives partial results on global max/min.

② Second derivative test + FOC:

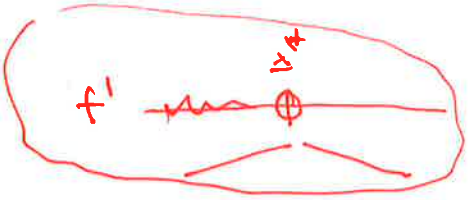
local data on f since these data are local → $f'_{x_1}(\underline{x}^*) = \dots = f'_{x_n}(\underline{x}^*) = 0$
and
$H(f)(\underline{x}^*)$ pos. defn. / neg. defn. / indefinite

local data
⇓
local classification

Ex: For a fn. in one variable, we have that

$$\left.\begin{array}{l}\text{one unique}\\ \text{stationary pt } x^*\\ \text{which is local max/min}\end{array}\right\} \Rightarrow x^* \text{ global max/min}$$

x^* local max:



x^* local min

This does not hold when f is fn. in several variables.

Ex: $f(x,y) = x^2y^3 + y^2 - 2y$

$f'_x = 2xy^3 = 0$ $\qquad$ $x=0$ or $y=0$

$f'_y = 3x^2y^2 + 2y - 2 = 0$ $\qquad$ $x=0$: $2y-2=0$, $y=1$ $\qquad$ $y=0$: $-2=0$ impossible

$\Downarrow$

$(0,1)$ is the only stationary pt.

$$H(f) = \begin{pmatrix} 2y^3 & 6xy^2 \\ 6xy^2 & 6x^2y+2 \end{pmatrix} \qquad H(f)(0,1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$D_1 = 2$
$D_2 = 4$
pos. def.
$\Downarrow$
$(0,1)$ local min

$f(0,1) = 0+1-2 = -1$ $\leftarrow$ not global min.
local min.

since for example
$f(3,-1) = -9+1+2 = -6 < -1$

Example with a unique stationary pt, which is local min, but not global min!

3 Convex / concave functions

Result:

If $H(f)(\underline{x})$ is positive semidefn. for all pts. $\underline{x}$, then f is convex.

If $H(f)(\underline{x})$ is negative semidefn for all pts. $\underline{x}$, then f is concave.

Result:

If f is convex, any stationary pt. is a global min,

If f is concave, any — " — is a global max