Lecture 4: Quadratic forms.

Def.: A function $Q(\vec{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j$ on $\mathbb{R}^n$.

Note: no linear / constant term, only the quadratic part of $c + \vec{p}^T\vec{x} + \underbrace{Q(\vec{x})}$.

"Why"?

What is the use of ax^2?

- $ax^2 + bx + c$, the "prototypical nonlinear function".
- also, prototypical concave / convex
  ↳ but: $Q(\vec{x})$ could be neither if $n > 1$, example: xy.
- under the hood of 2nd derivative tests and 2nd order cond's: a quadratic approximation.
- applications like, e.g.:

  Let $\vec{Y}$ be a random vector with $E[\vec{Y}] = 0$ and covariance matrix $\vec{A} = E[\vec{Y}\vec{Y}^T]$. The variance of $\vec{w}^T\vec{Y}$, ($\vec{w}$ nonrandom) is $\vec{w}^T\vec{A}\vec{w} = Q(\vec{w}) = \sum_i\sum_j a_{ij} w_i w_j$.

Matrix formulation:

$$Q(\vec{x}) = \vec{x}^T \vec{A} \vec{x} \quad \text{with } \vec{A} = (a_{ij})_{i,j=1}^{n}$$

$\vec{A}$ can be taken as <u>symmetric:</u>

$\to$ because $x_i x_j = x_j x_i$, we have

$$a_{ij} x_i x_j + a_{ji} x_j x_i = \tfrac{1}{2}(a_{ij} + a_{ji})(x_i x_j + x_j x_i)$$

$\to$ linear algebra: $Q(\vec{x})$ is a number,

So $\vec{x}^T \vec{A} \vec{x}$ = its own transpose

$$= \vec{x}^T \vec{A}^T (\vec{x}^T)^T$$

As $\vec{x}^T \vec{A} \vec{x} = \vec{x}^T \vec{A}^T \vec{x}$, both equal

$$\tfrac{1}{2} \vec{x}^T \underbrace{(\vec{A} + \vec{A}^T)}_{\text{symmetric.}} \vec{x}$$

The matrix tools to follow will require $\vec{A} = \vec{A}^T$.

- $Q(\vec{x}) = \vec{x}^T \vec{M} \vec{x}$ makes sense if $\vec{M}$ isn't symmetric, but <u>you are expected to</u> remember to rewrite as $\vec{x}^T \vec{A} \vec{x}$ by putting $\vec{A} = \frac{1}{2}(\vec{M} + \vec{M}^T)$.

Call the symmetric $\vec{A}$ the matrix associated with the function Q.

Definiteness:

For $n=1$, the function ax^2 is

- for $a>0$: strictly convex, and $ax^2>0$ except for $x=0$.
- for $a<0$: strictly concave and $ax^2<0$ except for $x=0$
- for $a=0$: concave and convex (not strictly!) and is zero on the entire line.

What about $n>1$?

Definition: Q, and its associated (symmetric!) matrix $\vec{A}$, will be called

- positive definite if $Q(\vec{x})>0 \ \forall \vec{x} \neq \vec{0}$
- positive semidefinite if $Q(\vec{x}) \geq 0$. all $\vec{x}$
- negative semidefinite if $Q(\vec{x}) \leq 0$, all $\vec{x}$
- negative definite if $Q(\vec{x})<0, \forall \vec{x} \neq \vec{0}$
- indefinite otherwise, i.e. if Q attains both values >0 and <0.

Ex.: $2xy$ is indefinite. $(x\ y)^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

Notes:

* Every $Q(\vec{x}) = \vec{x}^T \vec{M} \vec{x}$ is either pos. def., pos. semidef, neg def, neg. semidef or indefinite, but: these properties are not defined for a non-symmetric $\vec{M}$. You have to use $\frac{1}{2}(\vec{M} + \vec{M}^T)$ if $\vec{M}^T \neq \vec{M}$.

* Once we have defined (strict) concavity/Convexity, it shall turn out that

  Q pos. def $\Leftrightarrow$ Q strictly convex

  etc.

  and

  Q indefinite $\Leftrightarrow$ "Q nowhere convex and Q nowhere concave"

* Terminology like "nonnegative definite" etc., can be found in the Literature, but is less common.

* Some texts use, e.g. "$\vec{A} \geq B$" for "$\vec{A} - \vec{B}$ is positive semidefinite".

* "Positive definite function" means one of two distinct (but related) properties.

How to decide definiteness?

↝ "Analysis" vs "linear algebra"

↓

$Q(t\vec{x}) = t^2 Q(\vec{x})$ so it suffices to consider max/min $Q(\vec{x})$ s.t. $\vec{x}^T\vec{x} = 1$ (why?)

→ will be a seminar problem and will lead to linear algebra too!

Today: Criteria in terms of minors of $\vec{A}$ (← the symmetric, remember!)

Definitions: For a square matrix,
[we shall only have use for this for symmetric matrices.]

* a $(k \times k)$ principal minor is formed by deleting the "same-numbered rows and columns":

  If delete all columns $j \in J$, then we also delete all rows $j \in J$.

and

* a $k \times k$ leading principal minor deletes rows $j > k$ col's $j > k$ and retains the top-left $k \times k$ corner.

The book's notation:

D_r for the $r \times r$ leading principal minor
("r" is not rank!)

Δ_r for any of the $\frac{n!}{r!(n-r)!}$ principal minors (including D_r)

(when I say "the Δ_r are" it means <u>all</u>.)

Ex.:
$$\vec{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix} \quad \text{has} \quad D_2 = \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}$$

and six Δ_2's: D_2, $\begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix}$, $\begin{vmatrix} 1 & 4 \\ 4 & 7 \end{vmatrix}$,
$\begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix}$, $\begin{vmatrix} 3 & 5 \\ 5 & 7 \end{vmatrix}$ and $\begin{vmatrix} 5 & 6 \\ 6 & 7 \end{vmatrix}$.

Criteria: we have the following implications:

$$\left[\begin{array}{l} D_r > 0, \quad \text{all } r = 1, \dots n \\ \Updownarrow \\ \text{pos. def.} \\ \Downarrow \\ \text{pos. semidef} \\ \Updownarrow \\ \text{All } \Delta_r \geq 0, \quad (\text{all } r = 1, \dots n) \end{array}\right.$$

Since $\vec{A}$ neg. def $\Leftrightarrow$ $(-\vec{A})$ pos def. etc., we have the implications

$$\left[\begin{array}{l} (-1)^r D_r > 0, \quad \text{all } r = 1, \dots, n \\ \Updownarrow \\ \text{neg. def} \\ \Downarrow \\ \text{neg. semidef} \\ \Downarrow \\ \text{All } (-1)^r \Delta_r \geq 0, \quad \text{all } r = 1, \dots n. \end{array}\right.$$

(Why the $(-1)^r$?)

Ex. $\begin{pmatrix} 1 & 2 & \text{Whatever}_1^T \\ 2 & 3 & \\ \text{Whatever}_1 & & \text{whatever}_2 \end{pmatrix}$ is indefinite,

as $D_2 = \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = -1$

2×2: $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$

$\boxed{\Delta_2 = D_2}$

- Indefinite if and only if $ac - b^2 < 0$.
- If $ac - b^2 = 0$: neither pos. def nor neg. def,

  but: (Δ_1): $\begin{cases} \text{pos. semidef if } a \geq 0 \;\&\; c \geq 0 \\ \text{neg. semidef if } a \leq 0 \;\&\; c \leq 0. \end{cases}$

- If $ac - b^2 > 0$, then:

  $\begin{cases} \text{pos. def if } a > 0 \text{ and } c > 0 \\ \text{neg. def if } a < 0 \text{ and } c < 0 \end{cases}$

  If $a > 0 = c$ or $c > 0 = a$, we would have "pos. semidef but not pos def" – but if $ac = 0$, we cannot have $D_2 > 0$. Likewise for "neg. semidef but not neg. def", impossible if $D_2 > 0$.

Actually: If $\vec{A}$ pos. (neg.) semidef & $|\vec{A}| \neq 0$, then $\vec{A}$ pos. (neg.) definite!

$\rightsquigarrow$ Covariance matrices are <u>always</u> pos. semidef – just check invertibility!

To verify that the criteria work for 2×2, let us do that case thoroughly:

$Q(x,y) = ax^2 + 2bxy + cy^2$,

- Case $a=b=c=0$: Easy.
- Case $b=0$, $ac=0$: Semidef:
  $D_2=0$ and $Q = ax^2$ or $= cy^2$
  One $\Delta_1 = 0$, the other decides
  pos semidef or neg. semidef.
- Case $b \neq 0$, $ac=0$: Say, $a \neq 0 = c$
  (other case similar)

  $\underset{\neq 0}{a}x^2 + \underset{\neq 0}{b}xy = x\cdot(ax+by)$ Fix $x\neq 0$. Let y vary.
  Indefinite,
  as the criteria say: $D_2 < 0$
- Case $abc \neq 0$.

$$Q(x,y) = a\left[x^2 + 2\,\frac{b}{a}xy + \frac{c}{a}y^2\right]$$

$$= a\left[(x+\frac{b}{a}y)^2 + \left(\frac{c}{a} - \frac{b^2}{a^2}\right)y^2\right]$$

$$= a\cdot[\text{square}] + \frac{1}{a}(ac-b^2)\,y^2,$$

a and $\frac{1}{a}$ have same sign.

Case $abc \neq 0$ cont'd: assume $(x,y) \neq (0,0)$

$$Q(x,y) = a \cdot \left(x + \frac{b}{a}y\right)^2 + \frac{1}{a}\, y^2 \det \vec{A}.$$

note a and $\frac{1}{a}$ have same sign.

... if $\det \vec{A} < 0$, then $Q(x,0)$ and $Q(-\frac{b}{a}y, 0)$ have opposite signs. Indefinite.

... if $\det \vec{A} = 0$, then $Q(x,y)$

$$= a \cdot \underbrace{\left(x + \frac{b}{a}y\right)^2}_{\geq 0,\ \text{but } = 0 \text{ when } x = -\frac{b}{a}y}$$

semidef but not definite.

... if $\det \vec{A} > 0$: $y^2 \det \vec{A} \geq 0$

$\left(x + \frac{b}{a}y\right)^2 \geq 0$

not both zero except at $\vec{0}$.

For $n > 2$: Everything works by completing squares.

↳ 

these minors check $Q(0, x_2, x_3)$ etc.

3x3 example with parameter

For each $t \in \mathbb{R}$, decide the definiteness of

$$\bar{A}_t = \begin{pmatrix} 1 & 3 & 1 \\ 3 & t & -1 \\ 1 & -1 & t-4 \end{pmatrix}$$

[case $t=1$ already done, indefinite]

That is also: the definiteness of the function $Q(x,y,z) = (x\ y\ z)\,\bar{A}_t \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

- Observe: $\bar{A}_t$ symmetric!
- "At a glance": $D_1 = 1 > 0$, so $\bar{A}_t$ is not neg. semidef., any t.
  Also, $t<4 \Rightarrow$ indefiniteness (the (3,3) element).
- "Strategic considerations":
  Unless it so happens that $\bar{A}_t$ is never positive definite, we will have to calculate the signs of both D_2 and D_3

$D_2 = \begin{vmatrix} 1 & 3 \\ 3 & t \end{vmatrix} = t-9$. If $t<9$ then indefinite.

by last row

$$D_3 = \begin{vmatrix} 3 & 1 \\ t & -1 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 1 \\ 3 & -1 \end{vmatrix} + (t-4)\,D_2$$

$$= -t-7+t^2-13t+36 = \underline{t^2-14t+29}$$

Have: Because $\bar{A}_t$ not neg. semidef, it must be indefinite if $t<4$ [unnecessary, because next]

or if $t<9$

or if $t^2-14t+29<0$

Iff on the other hand $t>9$ and $t^2-14t+29>0$, then $D_1>0$, $D_2>0$, $D_3>0$

$\Rightarrow$ pos. def.

$\hookrightarrow$ $t^2-14t+29>0$ when?

$=0$: $t=7\pm\frac{1}{2}\sqrt{196-116}=7\pm2\sqrt{5}$

Convex parabola, negative between zeroes.

$\underbrace{7-2\sqrt{5}}_{<9}$ and $\underbrace{7+2\sqrt{5}}_{>9}$.

So: Pos. def $\Leftrightarrow$ $t>7+2\sqrt{5}$

If $t<7+2\sqrt{5}$: $D_1>0>D_3$ $\Rightarrow$ indefinite

- Remains: Case $t=7+2\sqrt{5}$

Either: calculate the remaining Δ_2's:

$\begin{vmatrix} 1 & 1 \\ 1 & t-4 \end{vmatrix} = t-5>0$ for $t=7+2\sqrt{5}$ | $\begin{vmatrix} t & -1 \\ -1 & t-4 \end{vmatrix} = t^2-4t-1$ For $t=7+2\sqrt{5}$, this is >0. $(\cdots 49+48\sqrt{5}+20-28-8\sqrt{5}-1)$

So: pos. semidef

Or: $Q(x,y,z)$ depends on t, continuously.

$$Q(x,y,z)\Big|_{t=7+2\sqrt{5}} = \lim_{t\searrow 7+2\sqrt{5}} Q(x,y,z)$$

$=\lim$ of $\underbrace{>0}_{\text{by pos. def}}$ which ≥ 0. $\Rightarrow$ pos semidef.

Eigenvalue characterization:

Let $\vec{A}^T = \vec{A}$ be $n \times n$. Then we have the following

facts:

- $\vec{A}$ has n lin. indep. eigenvectors.
- The char. pol. is $p(\lambda) = (\lambda_1 - \lambda) \cdot \ldots \cdot (\lambda_n - \lambda)$
  i.e. n real eigenvalues, counted with multiplicity.
- The eigenvectors $\vec{v}^{(i)}$ corr. to λ_i and $\vec{v}^{(j)}$ corr to λ_j, $i \neq j$:
  - are orthogonal $(\vec{v}^{(i)} \cdot \vec{v}^{(j)} = 0)$ if $\lambda_i \neq \lambda_j$
  - can be chosen orthogonal if $\lambda_i = \lambda_j$.

- $\vec{A}$ pos. def $\iff$ all $\lambda_i > 0$
  pos. semidef $\iff$ all $\lambda_i \geq 0$
  neg def $\iff$ all $\lambda_i < 0$
  neg semidef $\iff$ all $\lambda_i \leq 0$.

- $n=2$ for simplicity (generalizes!)
  Suppose $\vec{A}$ indefinite: $\underset{\vec{w}}{\lambda_2} > 0 > \underset{\vec{v}}{\lambda_1}$

  Then " Q convex along $\vec{w}$ and concave along $\vec{v}$"

  Ex: xy. $y=x$ yields x^2, convex
  $y=-x$ yields $-x^2$, concave. $\Big\}$ $\vec{0}$ is saddle!

Ex: the "previous" with $t = 1$. ($< 5+\sqrt{18}$, indef.)

Last time: eigenvalues $\underline{-4, -1, 4}$.

Ex: For every $t \in \mathbb{R}$, decide the definiteness of

$$\vec{A}_{t,n} = \vec{I}_n + t \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}}_{=: \vec{E}_n},$$

which has eigenvectors

$$\vec{e}^{(1)} - \vec{e}^{(2)},\ \vec{e}^{(1)} - \vec{e}^{(3)},\ \dots,\ \vec{e}^{(1)} - \vec{e}^{(n)} \text{ and } \vec{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Here, $\vec{e}_1 = (1,0,\dots,0)$, $\vec{e}_2 = (0,1,0,\dots)$, etc.

Calculate eigenvalues: $\vec{A}_{t,n}(\vec{e}^{(1)} - \vec{e}^{(i)})$

$$= \vec{e}^{(1)} - \vec{e}^{(i)} + t \cancel{\begin{pmatrix} \vec{1} \cdot (e^{(1)} - e^{(i)}) \\ \vdots \\ \vec{1} \cdot (\vec{e}^{(1)} - \vec{e}^{(i)}) \end{pmatrix}}$$

$$\lambda_i = \underset{>0}{1}, \quad i = 2, \dots$$

$$\vec{A}_{t,n} \vec{1} = \vec{1} + t \begin{pmatrix} \vec{1} \cdot \vec{1} \\ \vdots \\ \vec{1} \cdot \vec{1} \end{pmatrix} = \vec{1} + tn \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

so $\lambda_1 = 1 + tn$. If $t < -\frac{1}{n}$: <u>indefinite</u>.

If we can show that we have found all n lin. indep. eigenvectors, then: pos. def for $t > -\frac{1}{n}$

and: pos. semidef, but not pos. def, for $t = -\frac{1}{n}$.

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & \cdots & 0 \\ \vdots & 0 & -1 & & \vdots \\ & \vdots & 0 & \ddots & \vdots \\ & \vdots & \vdots & & 0 \\ 1 & 0 & \cdots & & -1 \end{vmatrix}$$

Exercise: add rows $2, 3, \dots, n$ to row 1 and then col's $2, 3, \dots, n$ to col. 1. Get something $\neq 0$.