Lecture 3: Optimization

① Definiteness of quadratic forms

Defn. A quadratic function in n variables $x_1, x_2, \ldots, x_n$ is a polynomial fn. of degree 2. A quadratic form is a function where all terms have degree 2.

Ex: $f(x_1,x_2,x_3,x_4) = x_1^2 - x_4^2 + 2x_2x_3$ quadratic form

Fact: Any quadratic form can be written $f(\underline{x}) = \underline{x}^T A \underline{x}$, with $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, for a unique $n \times n$ symmetric matrix A.

Ex: $f(\underline{x}) = x_1^2 - x_4^2 + 2x_2x_3$

$= \underline{x}^T A \underline{x}$ with $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

diagonal entries
– square coeff.
other entries
– cross terms

$\ldots + c_{ii} x_i^2 + \ldots \rightsquigarrow a_{ii} = c_{ii}$

$\ldots + c_{ij} x_i x_j + \ldots \; (i \neq j) \rightsquigarrow a_{ij} = a_{ji} = c_{ij}/2$

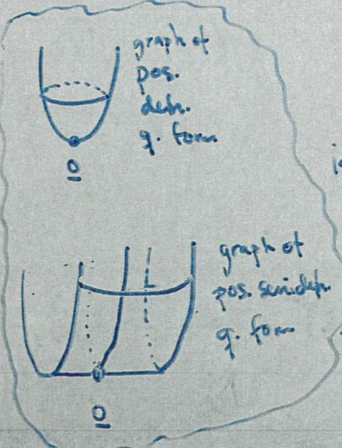
Defn: $f(\underline{x}) = \underline{x}^T A \underline{x}$ is called positive semidefn if $f(\underline{x}) \geq 0$ for all $\underline{x}$

negative –"– " $f(\underline{x}) \leq 0$ –"–

indefinite otherwise

is called positive defn. if $f(\underline{x}) > 0$ for all $\underline{x} \neq \underline{0}$

negative defn. " $f(\underline{x}) < 0$ –"–



Fact: $f(\underline{x}) = \underline{x}^T A \underline{x}$ quadr. form

$\lambda_1, \lambda_2, \dots, \lambda_n$ eigenvalues of A

f pos. semidefn. $\iff \lambda_i \geq 0$ for all i

f neg. semidefn. $\lambda_i \leq 0$ —"—

f indefinite otherwise

f positive defn. $\iff \lambda_i > 0$ —"—

f negative " $\lambda_i < 0$ —"—

Ex: $f = 2x^2 + 2xy + 2xz + 2y^2 + 2yz + 2z^2$
$= \underline{x}^T A \underline{x}$

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

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)\cdot\left((2-\lambda)^2 - 1^2\right) - 1\cdot\left((2-\lambda)-1\right) + 1\cdot\left(1-(2-\lambda)\right)$$

$$= (2-\lambda)(\lambda^2 - 4\lambda + 3) - (1-\lambda) + (\lambda - 1)$$

$$= (2-\lambda)(\lambda-1)(\lambda-3) + 2(\lambda-1)$$

$$= (\lambda-1)\left((2-\lambda)(\lambda-3)+2\right) = (\lambda-1)\cdot(-\lambda^2+5\lambda-4)$$

$$= -(\lambda-1)(\lambda^2-5\lambda+4) = -(\lambda-1)(\lambda-1)(\lambda-4) = 0$$

$$\lambda_1 = \lambda_2 = 1,\ \lambda_3 = 4 > 0$$

f pos. defn.

global min at $\underline{0}$

From theory: A symm. $\Rightarrow$ there is $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ and P with $P^T = P^{-1}$ s.t.

$$P^T A P = D$$

Change of var's

$\underline{x} = P\underline{u}$ i.e. $\underline{u} = P^T \cdot \underline{x}$ and $P^T A P = D \Rightarrow A = PDP^T$

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ old, $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ new

Expl: $f = \underline{x}^T A \underline{x}$

$= (P\underline{u})^T \cdot (PDP^T) \cdot (P\underline{u})$

$= \underline{u}^T \cancel{P^T P} D \cancel{P^T P} \underline{u}$

$= \underline{u}^T D \underline{u} = 1\cdot u^2 + 1\cdot v^2 + 4w^2$

pos. defn.

Note: $f(\underline{x}) = \underline{x}^T A \underline{x}$

quadr. form in n var's

$\lambda_1, \dots, \lambda_n$ eigenvalues of A, $n \times n$ symm. matrix

Assume: $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$

$$\boxed{\text{Then: } \lambda_1 \cdot \|\underline{x}\|^2 \le \underline{x}^T A \underline{x} \le \lambda_n \|\underline{x}\|^2 \text{ for all } \underline{x}}$$

Explanation:

$\{\underline{v}_1, \dots, \underline{v}_n\}$ orthonormal base of eigenvectors of A, with $A\underline{v}_i = \lambda_i \underline{v}_i$

$\Rightarrow$ Any vector $\underline{x}$ in $\mathbb{R}^n$ can be expressed as $\underline{x} = c_1 \cdot \underline{v}_1 + \dots + c_n \underline{v}_n$

Then $\|\underline{x}\|^2 = (c_1 \underline{v}_1 + \dots + c_n \underline{v}_n) \cdot (c_1 \underline{v}_1 + \dots + c_n \underline{v}_n)$

$= c_1^2 + \dots + c_n^2$

$\Rightarrow$

$$\begin{aligned}
\underline{x}^T A \underline{x} &= \underline{x}^T (A(c_1 \underline{v}_1 + \dots + c_n \underline{v}_n)) \\
&= \underline{x}^T (c_1 \lambda_1 \underline{v}_1 + \dots + c_n \lambda_n \underline{v}_n) \\
&= (c_1 \underline{v}_1 + \dots + c_n \underline{v}_n)^T \cdot (\lambda_1 c_1 \underline{v}_1 + \dots + \lambda_n c_n \underline{v}_n) \\
&= c_1^2 \lambda_1 + \dots + c_n^2 \lambda_n
\end{aligned}$$

$\Downarrow$

$$\begin{array}{ccccc}
\lambda_1 \|\underline{x}\|^2 & \le & \underline{x}^T A \underline{x} & \le & \lambda_n \cdot \|\underline{x}\|^2 \\
\| & & \| & & \| \\
\lambda_1 (c_1^2 + \dots + c_n^2) & \le & \lambda_1 c_1^2 + \dots + \lambda_n c_n^2 & \le & \lambda_n (c_1^2 + \dots + c_n^2)
\end{array}$$

Computing eigenvalues

- solving n-th order eqn's, hard even with computational support
- in general, you can get complex numbers $\leftarrow$ $\boxed{\text{Python: } j}$

  but not if A is symmetric

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

no real solutions

complex solutions: $\lambda = i$, $\lambda = -i$

Python: 0+j, 0-j

$$\boxed{\text{Complex no.: } a + bi \quad \left(\text{"} i = \sqrt{-1} \text{"}\right)}$$

② Unconstrained optimization

max/min $f(\underline{x}) = f(x_1, x_2, \dots, x_n)$

Stationary pts: $f'(\underline{x}) = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix} = \underline{0}$

Hessian: $H(f) = \begin{pmatrix} f''_{x_1 x_1} & \cdots & f''_{x_1 x_n} \\ \vdots & & \vdots \\ f''_{x_n x_1} & \cdots & f''_{x_n x_n} \end{pmatrix}$ $n \times n$ symm. matrix

Defn:

f convex if $H(f)(\underline{x})$ pos. semidefin. for all $\underline{x}$

f concave if " neg. semidefin. — " —

Result: If f is <u>convex</u>, any stationary pt is <u>minimum</u>.

— " — <u>concave</u> — " — <u>maximum</u>

Application: Unconstrained optimization of <u>quadratic fns</u>

Remark: Any quadratic fn. $f(\underline{x})$ in n variables can be written

$$f(\underline{x}) = \underbrace{\underline{x}^T A \underline{x}}_{\substack{\deg = 2 \\ \text{quadr.} \\ \text{form}}} + \underbrace{B \underline{x}}_{\substack{\deg = 1 \\ \text{lin.} \\ \text{form}}} + \underbrace{C}_{\substack{\deg = 0 \\ \text{const.}}}$$

where
- A symm. $n \times n$ matrix
- B $1 \times n$ matrix (row vector)
- C 1×1 matrix (const.)

Ex: $f(x,y,z) = \overbrace{x^2 + 4xy + 2y^2 + 6yz + 3z^2}^{\to A} + \underbrace{x - y + z}_{\to B} + 4$

$= \underline{x}^T A \underline{x} + B \underline{x} + C$

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 3 \\ 0 & 3 & 3 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix}$$

$$C = (4)$$

<u>Some useful formulas</u>

$f(\underline{x}) = \underline{x}^T A \underline{x} + B \underline{x} + C$:

i) $f'(\underline{x}) = 2A\underline{x} + B^T$

ii) $H(f) = 2A$

Note:
$H(f)$ has the same definiteness as A

Ex: (cont'd) $f'(\underline{x}) = 2A\underline{x} + B^T = 2 \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & 3 \\ 0 & 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $\quad A\underline{x} = -\frac{1}{2} B^T$.

$|A| = 1 \cdot (-3) - 2 \cdot 6$
$= -15 \neq 0 \Rightarrow A$ inv.

$\underline{x} = A^{-1} \cdot \left(-\frac{1}{2} B^T\right) = -\frac{1}{2} A^{-1} \cdot B^T$

One stationary pt.

<u>Definiteness of A:</u> $\operatorname{tr}(A) = 1 + 2 + 3 = 6 \Rightarrow A$ <u>indefinite</u>

$\det(A) = -15$

$\Rightarrow \underline{x} = -\frac{1}{2} A^{-1} B^T$ <u>saddle pt</u>

Some useful facts:

i) A symm. $n\times n$-matrix with eigenvalues $\lambda_1,\dots,\lambda_n$: $\begin{cases} tr(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n \\ det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n \end{cases}$

ii) A pos. defn $\iff$ A pos. semidefn and $|A| \neq 0$
" neg " " neg ——" ——

Fact: If A is any $m\times n$-matrix, then A^TA is pos. semidefn.

a) $\underset{n\times m}{A^T}\ \underset{m\times n}{A} \longrightarrow \underset{n\times n}{A^TA}$, $(A^TA)^T = A^T(A^T)^T = A^TA$

$\Rightarrow A^TA$ $n\times n$ symm. matrix

b) $\underline{x}^T(A^TA)\underline{x} = (\underline{x}^TA^T)\cdot(A\underline{x}) = (A\underline{x})^T\cdot(A\underline{x})$

$= \|A\underline{x}\|^2 \geq 0$ for all $\underline{x}$

$\Rightarrow A^TA$ is positive semidefn.

3 Constrained optimization:

$\max/\min f(\underline{x})$ when $\begin{cases} g_1(\underline{x}) = a_1 \\ g_2(\underline{x}) = a_2 \\ \vdots \\ g_m(\underline{x}) = a_m \end{cases}$

Method of Lagrange multipliers:

1) Form $L(\underline{x};\underline{\lambda}) = f(\underline{x}) - \lambda_1\cdot(g_1(\underline{x})-a_1) - \dots - \lambda_m(g_m(\underline{x})-a_m))$

2) Solve the eqn's $(*)$ $\boxed{\begin{matrix} L'(\underline{x}) = \underline{0} \\ g_1(\underline{x}) = a_1 \\ \vdots \\ g_m(\underline{x}) = a_m \end{matrix}}$. If the problem has a max/min, it is one of the solution obtained.

3) If $(\underline{x}^*;\underline{\lambda}^*)$ is a solution of $(*)$, then we have:

$\underline{x} \mapsto L(\underline{x};\underline{\lambda}^*)$ convex $\Rightarrow$ $\underline{x}^*$ min
—— " —— concave " max

We can apply matrix methods when $f, g_1, \dots, g_m$ are quadratic fn's.

Ex: $\min f(x,y,z) = x^2 + y^2 + z^2$ when $\begin{cases} x + y + z = 8 \\ 2x - y + 3z = 2 \end{cases}$

$f(\underline{x}) = \underline{x}^T I \underline{x} = \underline{x}^T \underline{x}$

$g(\underline{x}) = B_1 \underline{x}$

$g(\underline{x}) = B_2 \underline{x}$

$B_1 = (1 \ 1 \ 1)$

$B_2 = (2 \ {-1} \ 3)$

$L = \underline{x}^T \underline{x} - \lambda_1 (B_1 \underline{x} - 8) - \lambda_2 (B_2 \underline{x} - 2)$

$B_1 B_1^T = 3$

$B_1 B_2^T = 4 = B_2 B_1^T$

$B_2 B_2^T = 14$

$L'(\underline{x}) = 2\underline{x} - \lambda_1 B_1^T - \lambda_2 B_2^T = \underline{0}$

$B_1 \underline{x} = 8 \quad B_2 \underline{x} = 2$

$\Downarrow$

$\underline{x} = \frac{\lambda_1}{2} B_1^T + \frac{\lambda_2}{2} B_2^T$

$B_1 \underline{x} = 8: \quad \frac{\lambda_1}{2} B_1 B_1^T + \frac{\lambda_2}{2} B_1 B_2^T = 8 \quad \cdot 2$

$B_2 \underline{x} = 2: \quad \frac{\lambda_1}{2} B_2 B_1^T + \frac{\lambda_2}{2} B_2 B_2^T = 2 \quad \cdot 2$

$\Downarrow$

$3\lambda_1 + 4\lambda_2 = 8 \quad \cdot 4$

$4\lambda_1 + 14\lambda_2 = 2 \quad \cdot 3$

$12\lambda_1 + 16\lambda_2 = 32$

$12\lambda_1 + 42\lambda_2 = 6$

$-26\lambda_2 = 26$

$\underline{\lambda_2 = -1} \quad \lambda_1 = \frac{8 - 4\lambda_2}{3} = \underline{4}$

$\underline{x} = \frac{4}{2}\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{(-1)}{2}\begin{pmatrix}2\\-1\\3\end{pmatrix} = \begin{pmatrix}2\\2\\2\end{pmatrix} + \begin{pmatrix}-1\\1/2\\-3/2\end{pmatrix}$

$\begin{pmatrix}x\\y\\z\end{pmatrix} = \underline{\begin{pmatrix}1\\5/2\\1/2\end{pmatrix}}; \quad \begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix} = \underline{\begin{pmatrix}4\\-1\end{pmatrix}}$

$h = L(\underline{x}; 4, -1)$

$= \underline{x}^T I \underline{x} - 4(B_1 \underline{x} - 8) + 1(B_2 \underline{x} - 2)$

$H(h) = 2I$ pos. defin. $\Rightarrow h$ convex

$\Downarrow$

$\underline{x}^* = \begin{pmatrix}1\\5/2\\1/2\end{pmatrix}$ min pt.

$f(\underline{x}^*) = 1 + \frac{25}{4} + \frac{1}{4} = \frac{30}{4} = \frac{15}{2}$ min value

Ex: $\max 2x^2 + 2xy + 2xz + 2y^2 + 2yz + 2z^2$

when $x^2 + y^2 + z^2 = 12$

$f(\underline{x}) = \underline{x}^T A \underline{x} \qquad A = \begin{pmatrix}2 & 1 & 1\\1 & 2 & 1\\1 & 1 & 2\end{pmatrix}$

$g(\underline{x}) = \underline{x}^T I \underline{x}$

$= \underline{x}^T \underline{x}$

Eigenvalues:

$\lambda_1 = \lambda_2 = 1, \ \lambda_3 = 4$

$$\begin{array}{ccccc} & & f(\underline{x}) & & \\ & & \shortparallel & & \\ 1 \cdot \|\underline{x}\|^2 & \le & \underline{x}^T A \underline{x} & \le & 4 \cdot \|\underline{x}\|^2 \\ \shortparallel & & & & \shortparallel \\ 12 & & & & 48 \end{array}$$

$f_{max} = \underline{48}$ at $\underline{x} = 2\begin{pmatrix}1\\1\\1\end{pmatrix}$ with $\underline{\lambda = 4}$

4) Minimum variance portfolios

Construct portfolio by combining n securities

$\underline{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$ portfolio weights

$\underline{\mu} = (\mu_1\ \mu_2 \cdots \mu_n)$ vector of expected returns $\mu_i = E(R_i)$

$R_1, \dots, R_n$: return of the securities (stochastic var's)

$\Downarrow$

$\boxed{R = w_1 R_1 + \dots + w_n R_n}$ return of chosen portfolio

$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$ covariance matrix

$\sigma_{ii} = Var(R_i)$

$\sigma_{ij} = Cov(R_i, R_j)$ $i \neq j$

Assume:

① $w_1 + w_2 + \dots + w_n = 1$ (fully invested)

② $E(R) = \alpha$ (fixed expected return level α)

$\Downarrow$

Problem: $\min Var(R)$ when $\begin{cases} w_1 + w_2 + \dots + w_n = 1 \\ E(R) = \alpha \end{cases}$

Note: i) $f(\underline{w}) = Var(R) = \underline{w}^T \Sigma \underline{w}$, and $Var(R) \geq 0$ means that Σ pos. semidefin. symmetric $n \times n$-matrix

ii) $w_1 + \dots + w_n = \underline{1} \cdot \underline{w}$ where $\underline{1} = (1\ 1 \cdots 1)$

$E(R) = \underline{\mu} \cdot \underline{w}$ $\underline{\mu} = (\mu_1\ \mu_2 \cdots \mu_n)$

iii) If $|\Sigma| = 0$, then Σ has $\lambda = 0$ as an eigenvalue, and there is an eigenvector $\underline{w} \neq \underline{0}$ s.t. $\Sigma \underline{w} = \underline{0} \Rightarrow \underline{w}^T \Sigma \underline{w} = 0$, i.e. a nontrivial portfolio with $Var(R) = 0$. To avoid this situation, we assume

$|\Sigma| \neq 0 \iff \Sigma$ pos. defin.

Ex: min $\underline{w}^T \Sigma \underline{w}$ when $\underline{1}\cdot\underline{w} = 1$ (no assumption on expected return)

$$L = \underline{w}^T \Sigma \underline{w} - \lambda(\underline{1}\cdot\underline{w} - 1)$$

$$L'(\underline{w}) = \boxed{\begin{aligned} 2\Sigma\underline{w} - \lambda\cdot\underline{1}^T &= \underline{0} \\ \underline{1}\cdot\underline{w} &= 1 \end{aligned}}$$

$$\Sigma\underline{w} = \frac{\lambda}{2}\underline{1}^T$$

$$\underline{w} = \Sigma^{-1}\left(\frac{\lambda}{2}\underline{1}^T\right) = \frac{\lambda}{2}\Sigma^{-1}\underline{1}^T$$

$$\underline{1}\cdot\underline{w} = \underline{1}\left(\frac{\lambda}{2}\Sigma^{-1}\underline{1}^T\right) = 1$$

$$\frac{\lambda}{2}\left(\underline{1}\,\Sigma^{-1}\underline{1}^T\right) = 1$$

$$\longrightarrow \quad \lambda = \frac{2}{\underline{1}\cdot\Sigma^{-1}\underline{1}^T}$$

$$\underline{w} = \frac{\lambda}{2}\Sigma^{-1}\underline{1}^T = \frac{\Sigma^{-1}\cdot\underline{1}^T}{\underline{1}\cdot\Sigma^{-1}\cdot\underline{1}^T}$$

minimum variance portfolio

Note: Σ pos defn. $\Rightarrow \lambda_1, \dots, \lambda_n > 0$

Σ^{-1} exists and has eigenvalues $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n} > 0$

$\Rightarrow \Sigma^{-1}$ pos. defn. $\Rightarrow \underline{1}\,\Sigma^{-1}\underline{1}^T \neq 0$