Max/min – conceptual

$\vec{x}^*$ (global) maximum:

$$f(\vec{x}^*) \geq f(\vec{x}), \quad \text{all } \vec{x} \text{ in the domain of } f$$

strict : $>$ whenever $\vec{x} \neq \vec{x}^*$.

Min: analogous.

<u>Local</u> version:

To speak about "local", we need <u>open</u> <u>sets</u>.
In $\mathbb{R}^n$: open balls $\{\vec{x};\ \|\vec{x}-\vec{a}\| < r\}$, but could also use e.g. $\{\vec{x};\ \max_{i=1..n} |x_i - a_i| < r\}$:
important is, "some open $U \ni \vec{x}^*$" means "everywhere that is sufficiently close to $\vec{x}^*$"
(and "open" guarantees against $\vec{x}^*$ being a boundary point)

<u>Def</u>: $\vec{x}^*$ local max if there is some open $U \ni \vec{x}^*$ such that $f(\vec{x}^*) \geq f(\vec{x})$ for all $\vec{x} \in U$ for which f is defined.

Strict version, and loc. min: as you can think.

Max / min - open domains.

First-order condition for C^1 functions:

$$\nabla f(\vec{x}^*) = \vec{0} \quad (\text{really: } \vec{0}^T)$$

Have we covered anything that enables us to generalize?

- Well: if $\vec{0}$ is a supergradient for f at $\vec{x}^*$ .... max. Subgradient .... min.
- Allows us to decide for "known convex/concave" functions, but you already knew that $\|\vec{x}\|$ has strict global min at $\vec{x} = \vec{0}$.
- We have "uniqueness" results:

  If f is quasiconcave, then the set of global maxima is convex (possibly empty!).

  Strict quasiconvexity: unique (if exists)

Second-order cond's for C^2 functions

Suppose that $\nabla f(\vec{x}^*) = \vec{0}$.

We have the implications: ($\vec{H}$ = Hessian matrix)

$\vec{H}(\vec{x}^*)$ neg. def

$\Downarrow$

$\vec{H}(\vec{x})$ neg. def on some open $U \ni \vec{x}^*$

$\Downarrow$

.... same, except possibly at some isolated points (cf. x^4) $\Rightarrow$ f strictly concave on U.

$\Downarrow$

strict local max

$\vec{H}(\vec{x})$ neg. semidef on some open $U \ni \vec{x}^*$

$\Downarrow$

loc. max. & f concave on U.

[pos .... min ... likewise].

And:

$\vec{H}(\vec{x}^*)$ indefinite $\Rightarrow$ $\underbrace{\text{neither loc. max nor loc min}}_{=:\ \text{"saddle point"}}$.

So local 2^{nd} order cond's are just "localized" concavity / convexity tests; if the "some open $U \ni \vec{x}^*$" part can be replaced by "everywhere", then f is concave.

Lagrange's method:

$$\min / \max_{\vec{x} \in \mathbb{R}^n} f(\vec{x}) \quad \text{s.t} \quad \underbrace{\vec{g}(\vec{x})}_{m<n \text{ constraints}} = \vec{b}$$

$\vec{b}$: Const.

Lagrangian:

$$L(\vec{x}) = f(\vec{x}) - \vec{\lambda}^T \left(\vec{g}(\vec{x}) - \vec{b} \right)$$

Lagrange cond's:

$$\boxed{\begin{aligned} \nabla f(\vec{x}^*) &= \vec{\lambda}^T \frac{\partial \vec{g}}{\partial \vec{x}}(\vec{x}^*) \\ \vec{g}(\vec{x}^*) &= \vec{b} \end{aligned}}$$

Example:

constants

$$\min \ \vec{x}^T \vec{A} \vec{x} + \vec{\beta}^T \vec{x} \quad \text{s.t} \quad \vec{\Gamma} \vec{x} = \vec{c}$$

Cond's: $\vec{x}^T (\vec{A} + \vec{A}^T) + \beta^T = \lambda^T \vec{\Gamma}, \quad \vec{\Gamma} \vec{x} = \vec{c}$

Rewrite:

$$\boxed{\begin{pmatrix} \vec{0}_{m \times m} & \vec{\Gamma} \\ \vec{\Gamma}^T & \vec{M} \end{pmatrix} \begin{pmatrix} \vec{\lambda} \\ \vec{x} \end{pmatrix} = \begin{pmatrix} \vec{c} \\ \vec{\beta} \end{pmatrix} \quad \text{where } \vec{M} = -(\vec{A} + \vec{A}^T)}$$

$\begin{pmatrix} \vec{\lambda} \\ \vec{x} \end{pmatrix}$ notation for $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$ etc.

(<u>Not</u> curriculum; if $|\vec{M}| \neq 0$, the determinant is $|\vec{M}| \cdot \underbrace{|\vec{\Gamma} \vec{M}^{-1} \vec{\Gamma}^T|}_{\neq 0 \text{ if the leftmost } m \times m \text{ block of } \vec{\Gamma} \text{ is invertible.}} \cdot (-1)^{m+n}$.)

* Ineq. constraints:

$$\max f(\vec{x}) \quad \text{s.t.} \quad g_i(\vec{x}) \le b_i \quad i = 1, \ldots, m$$

$\uparrow$!

$$L = f(\vec{x}) - \vec{\lambda}^T(\vec{g}(x) - \vec{b})$$

Kuhn - Tucker cond's:

$$\boxed{\begin{array}{l} \nabla L(\vec{x}^*) = 0^T \\ \quad \lambda_i \ge 0 \quad \text{and, if } g_i(\vec{x}^*) < b_i: \ \lambda_i = 0. \end{array}}$$

Furthermore, the constraints hold.

So: for each i,

$$\underbrace{-\lambda_i}_{\le 0} \ (\underbrace{g_i(\vec{x}^*) - b_i}_{\le 0})$$

Product of two nonneg's, of which at least one is zero.

* Mixed constraints: some $g_i \le b_i$, some $h_i = c_i$

$$L = f(\vec{x}) - \vec{\lambda}^T(\vec{g}(x) - \vec{b}) - \vec{\mu}^T(\vec{h}(x) - \vec{c})$$

$$\boxed{\begin{array}{l} \nabla L(x^*) = 0^T \\ \quad \lambda_i \ge 0 \ (= 0 \text{ if } g_i < b_i) \\ \quad \vec{h}(x^*) = \vec{c} \end{array}}$$

* The Lagrange / K-T cond's are "close to necessary". Precise necessary cond's will be given after the teaching-free week.

* What about <u>sufficiency</u>?

Suppose that $\vec{x}^*$ satisfies the Lagrange / K-T cond's, producing multipliers $\vec{\lambda}$ (and $\vec{\mu}$ for that case)

<u>Fact:</u> If $\vec{x}^*$ maximizes L s.t. the constraints then $\vec{x}^*$ maximizes f ——— s.t. ———.
(Here $\vec{\lambda}$ (& $\vec{\mu}$) are these given numbers.
For the Lagrange case "min" works likewise.)

Why? Suppose $L(\vec{x}^*) \geq L(\vec{x})$. Then

$$0 \leq L(\vec{x}^*) - L(\vec{x}) = f(\vec{x}^*) - f(\vec{x}) - \sum_{j;\lambda_j>0} \lambda_j \overbrace{(\underbrace{g_j(x^*)}_{=b_j} - g_j(x))}^{\geq 0}$$

So $f(\vec{x}^*) \geq f(\vec{x}) + \sum_{j;\lambda_j>0} \lambda_j \cdot [\text{nonneg}] \geq f(\vec{x})$.

Example: Fix $\vec{v}$ s.t. $\|\vec{v}\| = 1$.

max/min $\quad \vec{x}^T \vec{v}$ s.t. $\vec{x}^T \vec{x} = 1$.

$L = \vec{v}^T x - 2\lambda(\vec{x}^T \vec{x} - 1)$

Cond's: $\left.\begin{array}{l} \vec{v}^T = \lambda \vec{x}^T \\ \vec{x}^T \vec{x} = 1 \end{array}\right\}$ so $\underbrace{\|2\lambda \vec{x}\|^2}_{4\lambda^2 \vec{x}^T \vec{x}} = \|\vec{v}\|^2 = 1$ so $\lambda = \pm 1/2$

Two points: $\vec{x} = \pm \vec{v}$ with $\lambda = \pm 1/2$ respectively

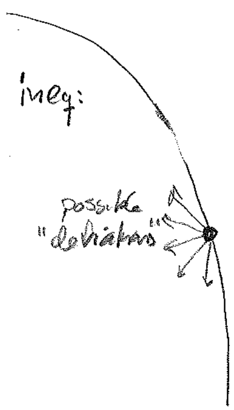
$\vec{x} = \vec{v}$: $\lambda = 1/2$ so L is concave. <u>max</u>

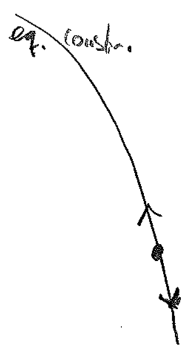
$\vec{x} = -\vec{v}$: $\lambda = -1/2$ so L is convex. <u>min</u>.

* Sufficient for $\vec{x}^*$ to maximize L
  (recall that $\vec{x}^*$ is a stationary point for L)
  is that $L(\vec{x})$ is concave in $\vec{x}$ (when the numbers are inserted for the multipliers)

* We can consider local versions:
  If the Hessian of L at $\vec{x}^*$ is neg. def., then $\vec{x}^*$ is a local max

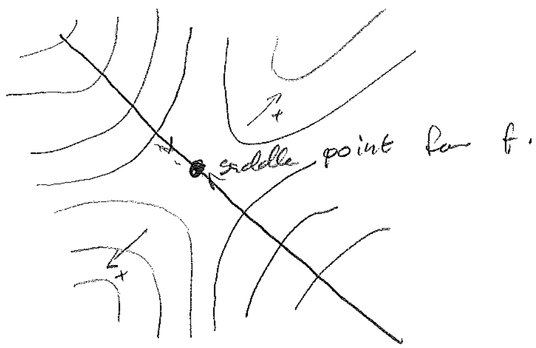
* This is not so powerful if the constraints are equalities. Sketch:





Can only deviate "along the constraint"

E.g., level curves



Concavity fails for this problem: the point is highest on the path, but any step in the direction would improve if it were admissible

Local 2^{nd} order cond's for equality constraints

$\max / \min \; f \quad s.t \quad \vec{G} = \vec{b}$

!!

Suppose $\vec{x}^*$ satisfies the Lagrange cond's.

Let $\vec{A}$ = the Hessian of f at $\vec{x}^*$

$\vec{B} = \frac{\partial \vec{G}}{\partial \vec{x}}(\vec{x}^*)$.

If $\vec{A}$ is neg. def subject to $\vec{B}\vec{v} = \vec{0}$

resp pos. def ———— , . ————

then $\vec{x}^*$ is a local max (resp. local min) subject to the constraint $\vec{G}(\vec{x}) = \vec{b}$.

(And: we have determinant criteria!)

What about ineq. constraints?

Suppose that $\vec{x}^*$ satisfies the Kuhn-Tucker cond's; can we

- → impose equality for those constraints 'active' at $\vec{x}^*$
  (i.e: if $g_j(\vec{x}^*) = b_j$, impose $g_j(x) = b_j$)
- → disregard the others
- → then we have a Lagrange problem to test with the previous cond's ?

Answer: nearly so – it is ok if $\lambda_j > 0$ for all active constraints.
(I.e: none have $\lambda_j = g(\vec{x}^*) - b = 0$.)

[Next example: not the local cond's!]

Example/application: The mutual fund theorem

a.k.a two-fund (monetary) separation

Model for risk-averse agent with

- one safe investment opportunity,
- n risky, excess returns distributed
  (above risk-free rate, "$r - r_f$" in CAPM lingo)

  with mean $= \vec{m}$, covariance matrix $\vec{A}$
- $\vec{A}$ assumed invertible.
  (Were it not, there would either be redundant opportunities – leave out & respecify! – or there would be a free lunch...)

Model: whatever the expected excess return $\vec{\eta}^T\vec{x}$ is, the agent minimizes variance $\vec{x}^T\vec{A}\vec{x}$

$$\min\ \vec{x}^T\vec{A}\vec{x} \quad \text{subject to} \quad \vec{\eta}^T\vec{x} = d$$

[Exercise: what would be the interpretation if the constraint were $\vec{\eta}^T\vec{x} \geq d$? What difference would it make?]

Rewrite to $\max\ (-\vec{x}^T\vec{A}\vec{x})$ s.t. $d - \vec{\eta}^T\vec{x} \begin{cases} = 0 \text{ or} \\ \leq 0 \end{cases}$;

$$L(\vec{x}) = -\vec{x}^T\vec{A}\vec{x} - \lambda(d - \vec{\eta}^T\vec{x}).$$

Stationarity: $2\vec{x}^T\vec{A} = \lambda\vec{\eta}^T$

so $\vec{x} = \underbrace{\vec{A}^{T-1}}_{=\vec{A}^{-1},\ \text{why?}}\vec{\eta}\ \frac{\lambda}{2}$;

$\vec{x} = \frac{\lambda}{2} \vec{A}^{-1} \vec{\eta}$ means that all agents choose the same – up to scaling – risky portfolio (= the "market portfolio" in CAPM lingo) ("Two"-fund separation: the other is the riskless.)

Once calculated out, we know (because $\max \underbrace{-\vec{x}^T \vec{A} \vec{x}}_{\text{concave}}$ s.t. [linear] is a concave program) that it solves the problem:

$d = \vec{\eta}^T \vec{x} = \frac{\lambda}{2} \underbrace{\vec{\eta}^T \vec{A}^{-1} \vec{\eta}}_{1 \times 1}$ yields

$$\frac{\lambda}{2} = \frac{d}{\vec{\eta}^T \vec{A}^{-1} \vec{\eta}} \quad \text{and} \quad \underline{\underline{\vec{x} = \frac{1}{\vec{\eta}^T \vec{A}^{-1} \vec{\eta}} \vec{A}^{-1} \vec{\eta}}}$$

"Discussion": an agent that is not risk-averse, might rather wish to maximize excess return subject to risk:

$$\max \vec{\eta}^T \vec{x} \quad \text{s.t.} \quad \vec{x}^T \vec{A} \vec{x} = R^2 .$$

More general, but has some technical quirks; <u>which</u>?