Optimization

Day 1: intuition
Day 2: proofs and definitions
Day 3: Applications

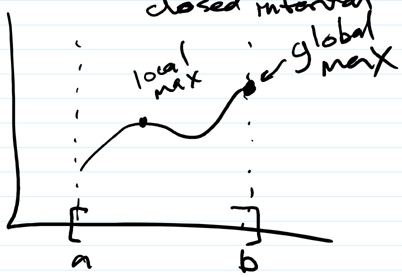
Unconstrained optimization

$$\max_{x_1,\dots,x_n} f(x_1, x_2, \dots, x_n)$$

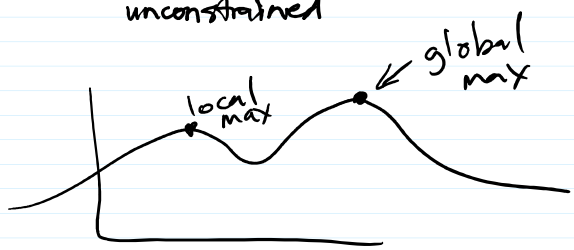
No border/boundary
→ every max/min occurs at a critical point, where $f'(x)=0$

or $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \cdots = \frac{\partial f}{\partial x_n} = 0$

Ex
constrained by the closed interval



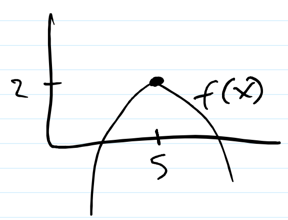
unconstrained



Solution: 1) Find all critical points
2) Take the largest one

Sometimes useful to check 2nd order conditions if unsure if it is a local min or max.

$f(x) = -(x-5)^2 + 2$



$$\max_{x} -(x-5)^2 + 2$$

Step 1): 1st order conditions

$f'(x) = 0$

$-2(x-5) = 0$

$\Rightarrow \boxed{x^* = 5}$ x^* is the maximizer

Check 2nd order conditions

$f''(x) < 0$ for a max

$f''(x) = -2 \rightarrow$ it is a local max

$\rightarrow$ since there is only one critical point, this is a global max.

Ex Least squares w/ no constant

$$y_i = \beta x_i + \underbrace{\varepsilon_i}_{\text{error}}$$

$$\boxed{\min_{\beta} \sum_{i=1}^{N} (y_i - \beta x_i)^2}$$

FOC's: $\sum_{i=1}^{N} 2(y_i - \beta x_i)\cdot(-x_i) = 0$

$$\sum_{i=1}^{N} (y_i x_i) - \beta \sum_{i=1}^{N} x_i^2 = 0$$

$$\Rightarrow \sum y_i x_i = \beta \sum x_i^2$$

$$\Rightarrow \beta^* = \frac{\sum y_i x_i}{\sum x_i^2}$$

S.O.C: $\frac{d^2}{d\beta^2}\left(\sum_{i=1}^{N}(y_i - \beta x_i)^2\right) > 0$

2nd deriv: $\frac{d}{d\beta}\left(-2\sum_{i=1}^{N}(y_i x_i - \beta x_i^2)\right)$

$$= -2\cdot\left[\underbrace{\frac{d}{d\beta}\sum y_i x_i}_{0} - \frac{d}{d\beta}\sum \beta x_i^2\right]$$

$$= -2\cdot(-\sum x_i^2) = 2\underbrace{\sum x_i^2}_{\text{all these are positive}} > 0$$

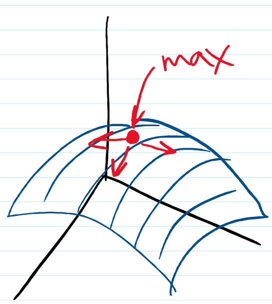
Multivariate

Now: FOC: $\frac{\partial f}{\partial x_1} = 0$

$$\frac{\partial f}{\partial x_2} = 0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = 0$$



SOC: Min: $H(x^*)$ positive definite [sufficient]

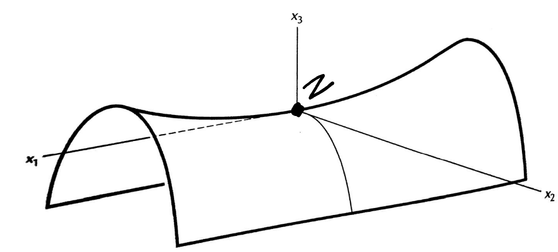
positive semidefinite [necessary]

SOC: Min: $H(x^*)$ positive definite [sufficient]
positive semidefinite [necessary]

Max: $H(x^*)$ negative definite [sufficient]
negative semidefinite [necessary]

→ try the sufficient conditions first.

Saddle Point: 2nd order necessary conditions fail at Z because it not a max or min.



The graph of the indefinite form $Q_3(x_1, x_2) = x_1^2 - x_2^2$.

Source: Simon & Blume page 378

Ex | $f(x,y,z) = e^{x-y} + e^{y-x} + e^{x^2} + z^2$

FOC's: (1) $f_1 = e^{x-y} - e^{y-x} + e^{x^2} \cdot 2x = 0$ ← set

(2) $f_2 = -e^{x-y} + e^{y-x} = 0$ ← set

(3) $f_3 = 2z = 0$ ← set

$\boxed{z=0}$ ✓

(2) $\Rightarrow e^{y-x} = e^{x-y}$

$$\frac{e^y}{e^x} = \frac{e^x}{e^y}$$

$$(e^y)^2 = (e^x)^2$$

(4) $\Rightarrow \boxed{x=y}$ since exp. fn is one-to-one

plug (4) into (1) $e^{x-y} - e^{y-x} + e^{x^2} 2x = 0$

$$\Rightarrow \underbrace{\cancel{e^0} - \cancel{e^0}}_{0} + e^{x^2} \cdot 2x = 0$$

$$\Rightarrow e^{x^2} \cdot 2x = 0 \qquad e^x \text{ is always} > 0$$

so $\Rightarrow \boxed{\begin{matrix} x^* = 0 \\ y^* = 0 \end{matrix}}$

and $\boxed{z^* = 0}$

Check SOC

$$H = \begin{bmatrix} e^{x-y} + e^{y-x} + e^{x^2} \cdot 2 + 2x e^{x^2} \cdot 2x & & \\ -e^{x-y} - e^{y-x} & e^{x-y} + e^{y-x} & \\ 0 & 0 & 2 \end{bmatrix}$$

$$f_1 = e^{x-y} - e^{y-x} + e^{x^2} \cdot 2x$$

$$f_2 = -e^{x-y} + e^{y-x} \qquad f_{21} = -e^{x-y} - e^{y-x}$$

$$f_3 = 2z \qquad f_{31} = 0$$

$$f_{32} = 0$$

SOC:

$$H(0,0,0) = \begin{bmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

condition for positive definiteness:
all leading principal minors of the Hessian must be positive.

i,jth minor is the determinant of the matrix formed by removing row i and col j.

A principal minor is the Det. of the matrix formed by removing any K rows and the K corresponding columns.

A leading principal minor is a principal minor that removes only rows/columns on the bottom right of the matrix.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad 2,3 \text{ minor is: } \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix}$$

Principal minors: $\begin{vmatrix} 1 & 3 \\ \end{vmatrix}$ remove: R2, C2

Principal minors: $\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix}$ remove: R2, C2

$\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix}$ R1, C1

$|5|$ R1, C1 and R3, C3

Leading principal minors:

$|1|$ remove R2, C2 and R3, C3

$\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}$ remove R3, C3

$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$

minors are the determinants of the "leading principal submatrices"

SOC:

$$H(0,0,0) = \begin{bmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

condition for positive definiteness:
all leading principal minors of the Hessian must be positive.

$|4| > 0$ ✓

$$\begin{vmatrix} 4 & -2 \\ -2 & 2 \end{vmatrix} = 4 \cdot 2 - 4 = 4 > 0$$

$$\begin{vmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}$$

how to find determinant? $(= 8 > 0)$ (see below)

Method #1

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 1 \cdot 6 \cdot 8 - 2 \cdot 4 \cdot 9 - 3 \cdot 5 \cdot 7$$

$$4 \cdot 2 \cdot 2 - (-2)(-2) \cdot 2 = 16 - 8 = 8$$

Method #2 Cofactors: expand along 3rd row

$$\text{Det} = 0 \cdot C_{31} + 0 \cdot C_{32} + 2 \cdot C_{33}$$

$$= 2 \cdot \underbrace{\begin{vmatrix} 4 & -2 \\ -2 & 2 \end{vmatrix}}_{=4} = 8$$

Method #3 Gaussian elim.

$$\begin{pmatrix} 4 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \overset{\frac{1}{2}R_1 + R_2}{\sim} \begin{pmatrix} 4 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

doesn't change determinant

AND this Det is just:

$4 \cdot 1 \cdot 2 = \underline{8}$

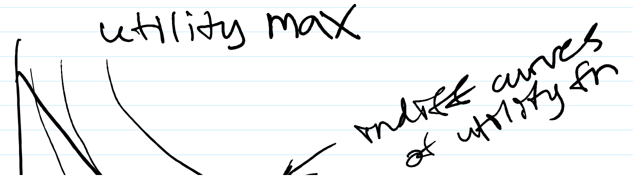
In conclusion: all leading principal minors of the Hessian are positive. Therefore this is a <u>minimum</u>.

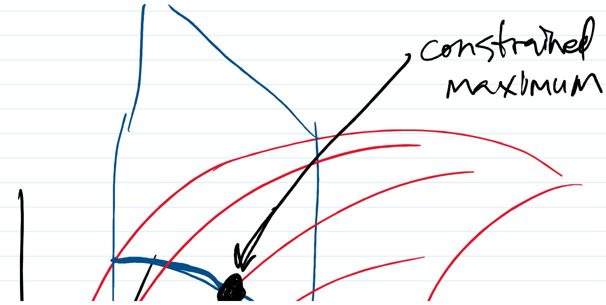
So $(x, y, z) = (0, 0, 0)$ is the <u>minimizer</u> point.

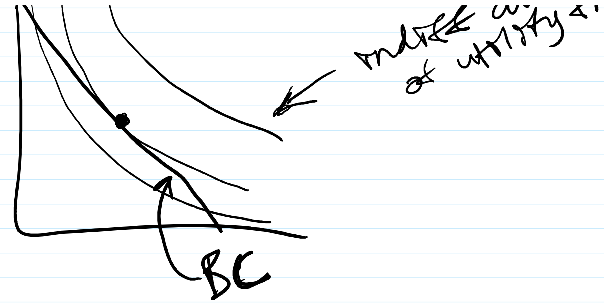
Constrained Optimization

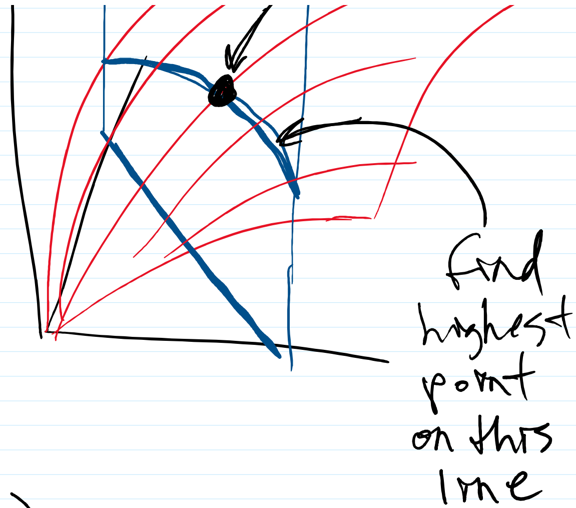
i. equality constraints

utility max





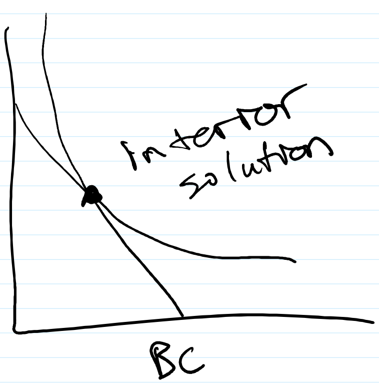


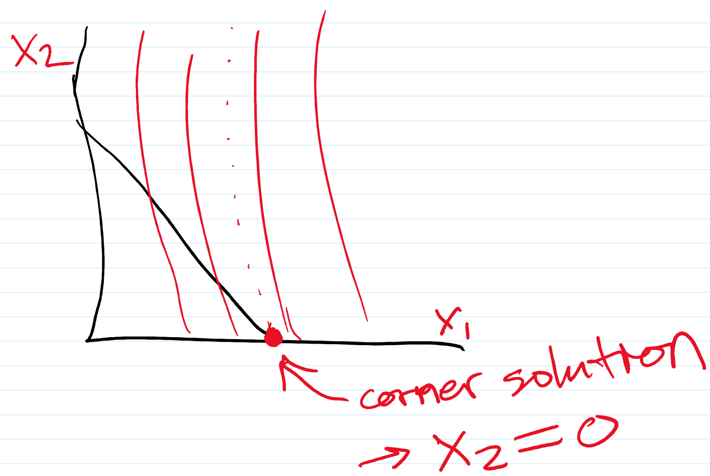


Two types of solutions

1. interior solution (not on the boundary)
2. corner solution (on the boundary)

usually in econ, the boundary
B : $x_1 \geq 0, x_2 \geq 0, \ldots$





Interior solutions

Apply the logic of unconstrained optimization but restrict the choice variable to lie on some line or plane (or some surface).

2 approaches:

1) Lagrangian

objective function

Lagrange multiplier

constraint

2 approaches:

1) Lagrangian
2) Gradient

1) Lagrangian: $$\mathcal{L} = \overset{\text{objective function}}{f(x)} + \overset{\text{Lagrange multiplier}}{\lambda}\left(\overbrace{c-g(x)}^{\text{constraint}}\right)$$

Used to solve: $\max\limits_{x} f(x)$ subject to: $c - g(x) = 0$

FOC:
$$\frac{\partial \mathcal{L}}{\partial x_1} = f_{x_1} - \lambda g_{x_1} = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = f_{x_2} - \lambda g_{x_2} = 0$$
$$\vdots$$
$$\frac{\partial \mathcal{L}}{\partial x_n} = f_{x_n} - \lambda g_{x_n} = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = c - g(x) = 0 \quad \leftarrow \text{constraint}$$

---

2) Gradient: $f_{x_1} - \lambda g_{x_1} = 0$ or

$$\left.\begin{aligned} f_{x_1} &= \lambda g_{x_1} \\ f_{x_2} &= \lambda g_{x_2} \\ &\vdots \\ f_{x_n} &= \lambda g_{x_n} \end{aligned}\right] \quad \overset{\text{gradient}}{\nabla} f = \lambda \nabla g$$

same for each equation

$$\Rightarrow \frac{f_{x_1}}{g_{x_1}} = \frac{f_{x_2}}{g_{x_2}} = \cdots = \frac{f_{x_n}}{g_{x_n}} = \lambda$$

Total differentials

$df = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$ → linear approximation of Δf as you move in the direction of $(dx_1, dx_2, \dots, dx_n)$

FOC in unconstrained opt:

$$df = 0 \text{ for all } (dx_1, dx_2, \dots, dx_n)$$

$df = \nabla f \cdot \vec{dx}$ only for all $\vec{dx}$ if $f_1 = f_2 = \cdots = f_n = 0$

if any $f_i \neq 0$, then

$$\nabla f \cdot (0\ 0\ 0 \cdots \underset{\substack{\uparrow \\ dx_i}}{a}\ 0\ 0\ 0)$$

$$= f_i \cdot a \neq 0 \quad \text{so } df \neq 0.$$

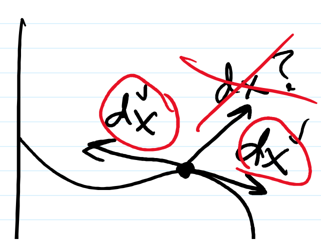
Constrained version:

$df = \nabla f \cdot dx = 0$ whenever $\underbrace{\nabla g \cdot dx = 0}$

when you move in the direction of dx, then $g(x)$ stays constant (very close to x)

so $g(x) = c$ will still be satisfied.

Note: not all dx will work.



if $\nabla f = \lambda \nabla g$

then when $\nabla g \cdot dx = 0$

we have $\nabla f \cdot dx = \lambda \nabla g \cdot dx$

$$= \lambda \cdot 0$$

$$= 0$$

So this implies that

So this implies that

$df = 0$ as we move along the constraint

→ that's exactly what we need for a critical point.

So an equivalent FOC for the constrained problem is:

$$\begin{cases} \nabla f = \lambda \nabla g \\ c - g(x) = 0 \end{cases}$$

Ex| $\max\limits_{x_1, x_2} -ax_1^2 - bx_2^2$

s.t. $x_1 + x_2 = 1 \longrightarrow \underbrace{g(x_1, x_2)}_{x_1+x_2} = \underset{\uparrow 1}{c}$

FOC's: $\nabla f = \lambda \nabla g$

$$\begin{cases} -2ax_1 = \lambda \cdot 1 & (1) \\ -2bx_2 = \lambda \cdot 1 & (2) \\ x_1 + x_2 = 1 & (3) \end{cases}$$

$(3) \implies \boxed{x_2 = 1 - x_1}$

Now take the ratio of (1) and (2) this will eliminate λ.

$$\frac{-2ax_1}{-2bx_2} = 1$$

$$\Rightarrow ax_1 = bx_2$$

Now plug in $x_2 = 1 - x_1$

$$\Rightarrow ax_1 = b(1-x_1) = b - bx_1$$

$$\Rightarrow (a+b)x_1 = b$$

$$\boxed{x_1 = \frac{b}{a+b}}$$

$$\boxed{x_2 = 1 - x_1 = \frac{a}{a+b}}$$

Multiple constraints

$$\mathcal{L} = f(x) + \lambda_1(c_1 - g^1(x)) + \lambda_2(c_2 - g^2(x))$$

Lagrange's theorem: there is a unique set of $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \cdots + \lambda_m \nabla g_m$$

true whenever $\nabla g_1, \ldots, \nabla g_m$ are linearly indep.

Interpretation of λ:

"Shadow price"

Interpretation of λ:

"shadow price"

How much will the objective fn increase when we relax the constraint?

Ex| consumer problem:

$$\max_{x,y} U(x,y) \quad \text{s.t. } P_x X + P_y Y = W$$

λ is how much the optimal U changes when W (wealth) increases.

optimal U — "value function" → optimal value of objective fn

$U = f(X,Y)$

$V = f(P_x, P_y, W)$ — "indirect utility fn" is a value fn.

$$\lambda = \frac{\partial V}{\partial c} \qquad \text{where } g(x) = c$$

consumer problem: utility value of additional money

λ "shadow price": intertemporal tradeoffs. How much am I willing to pay now for more consumption in period $t+n$?

Ex| Producer theory: how much is the producer willing to pay to relax a constraint? "shadow price"