OPTIMIZATION

Dr. Mohammad Al-Shannag

Optimization 28/1/2018

Introduction -

Optimization :

- Finding the best situations
- searching for either maximum or minimum.

اقتران هدف

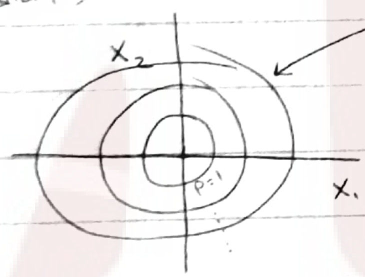
Max objective function] طريقتين
Min (o.f) symmetry.

Example: $P(x) \equiv$ cost, Profit, conversion, yield, ..., etc.

Ex. ~~Max p(x)~~ min. $p(x_1, x_2)$ intersection between equations → graphically can b

هذا اقتران لما ممكن يكون اكثر
لحصول المساواة

$x_1, x_2, \ldots, x_n$ → Decision variables.



isocontours of p : the line of the same value of the fun

e.g $p(x_1, x_2) = x_1^2 + x_2^2$ معادلة دائرة.

$x_1^* = 0 \quad x_2^* = 0 \rightarrow P(x_1, x_2)$

How $P=1$ $\quad x_1 =$ some value find x_2

Objective of the course :

(A) how to formulate the required optimization problem.

optimization problem - objective function $p(x)$ (make formulation).

- the corresponding constrains

(B) how to ~~find~~ solve the problem to obtain the optimum

(C) how to analyze the results (sensitivity of the results).

constants from physical properties, assumptions assumed.

so if these constants has changed.

Scanned by CamScanner

↑variables ↑complexity.

* Fields of applications

- economics, bussinuss, and management
- engineering ~~(operation)~~ design ≺ area / diameter / size
- engineering operation → optimum operating conditions.
- Physical Science → new impirical relations (curve fitting
  e.g least square methode,
  best constants in relation → optimization.

How to optimize?

Ⓐ model-based approach ← we are here
Ⓑ experimental approach (empirical)

Ⓐ Model-based approach:

- require mathmatical model (Disadvantage)
- Process (real situation) is not necessary exist.
- experiments are not needed for optimization. (experiments maybe required in modeling stage)
- optimization Phase is fast.
- ~~optimization~~
- model develepement phase ~~ff~~ maybe slow.
- evaluate different ~~set~~ sconarions easly.

Ⓑ Empirical optimization

- with or without mathmatical model (maybe semi-impirical)
- Process must be exist
- experiments are required for optimization.
- measure the objective function. قياس عملي او حسبان من مقادير
- optimization phase is slow (depends on response)
- no delay for modeling لايوجد داع للتفكير ب model.

Scanned by CamScanner

- Choose initial guess from the physical situation.
- Challenge ⌈ stepsize
  direction.

Approach of empirical optimization :

هو صح إنه
تثبت كل المتغيرات
وتغير واحد منهم
واحد عند ٤٢٥
مخطط عاكي أحسن
من ٣٠٠

- Perform (عمل تجارب) near the current operating point.
- Find the direction which improves the value of $p(x)$ (↑assumptions ↓accuracy)
- move along that direction
- repeat until converge to the optimum

Approach of model optimization :

- Formulate the optimization problem.
- start from some initial guess point. (if not hypothetical, it is the operation point).
- Find the direction which improves the $p(x)$ value using the model
- move along that direction (optimum stepsize) to optimize the target
- repeat until converge.

General optimization of the Formulation problem

max. or min. $P(\underline{X}) = \dots$

in book in bold, to indicate a vector variable, it can be single variable.

Subject to: (s.t) ← constrains $\underline{f}(\underline{x}) = \underline{0}$ "equality constrains"

$\underline{g}(\underline{x}) \geq \underline{0}$ "inequality constrains"

$\underline{x}_{Min} \leq \underline{x} \leq \underline{x}_{Max}$ "Variable bounds"

optimum
Put star

$P(\underline{x})$: scalar (1 objective function)

$\underline{x}$: column vector of $\overset{n}{?}$ decision variables $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$\underline{f}$: column vector of equations of $\overset{\text{مش}\ n}{m_1}$ dimension $\cancel{f_{:}}$

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{m_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$f_1(x_1, x_2, \dots, x_n) = 0$

$f_2(x_1, x_2, \dots, x_n) = 0$

$\vdots$

$f_{m_1}(x_1, x_2, \dots, x_n) = 0$

g : " " " m_2 inequality constrains $\geq \quad \times -1 \rightarrow \leq$ بعد النقل عشان يساوي صفر.

if $n = 3$ $m_1 = 3$, it doesn't need optimization. $\#DOF = 0$

if in optimization $\#$ Variables $>$ $\#$ constrains.

30/1/2019

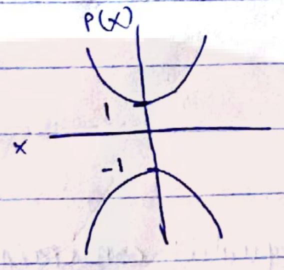
Optimization

$\Rightarrow$ max $p(x) \equiv$ min $p(x)$

e.g $p(x) = 1 + x^2$

$x^* = 0$ $p(x^*) = 1$

it is min. $p(x)$



max $g(x) = -p(x) = -(1+x^2)$

So we are not interested of min or max, Just multiply by -1

- If the optimization problem has no (A) $\underline{F}(x) = 0$, $\underline{g}(\underline{x}) \geq 0$ and without variable bound

It is called unconstrained optimization problem

e.g least square method

(B) other wise, it is constrained.

$\rightarrow$ Examples on equality constrains ($\underline{F}(\underline{x}) = 0$) sources

e.g - Physical laws as conservation principles of mass, energy, momentum, ... etc

ideal gas law, thermodynamics law, rxn, $\rightarrow$ from mole conservation.

- impirical correlations such as $Nu = a\, Re^b\, Pr^c$

some limit $\downarrow$ variable bound.

$\rightarrow$ Inequality constrains sources ~~(from~~ not equal or $\leq$

$\rightarrow$ e.g - Physical limitations ~~such as~~

combination of variables ~~Vapo ration~~

Ex: Flodding ~~distillation~~ condition in the distillation column

(equipment design). $\frac{\dot{V}}{\dot{L}} \leq c$ (some certain value).

$\rightarrow$ Variable bounds sources

- Physical limitations on the decision variables ($\underline{x}$):

e.g (conversion, Temp, flow rate,)

($\underline{x}$) : T_{Max}, P_{Max}, $Flow_{max}$, X_A, ...

$\Rightarrow$ Linear programming (LP) problem

LP problem is an optimization Problem with linear functions of $p(\underline{x})$, $\underline{g}(\underline{x})$, (may be more than one) and $\underline{F}(\underline{x})$. It always has $\underline{g}(x)$

$P(\underline{x}) = c_1 x_1 + c_2 x_2 + c_3 x_3 + \ldots$

optimum of $2x$ is ∞ & $-\infty$, for this reason we need $g(x)$

⇒ Non-linear programming (NLP)

Either $p(\underline{x})$, $\underline{f}(x)$, or $\underline{g}(x)$ is non linear

⇒ Quadratic programming problem ~~(QPP)~~ (QP)

In which the $p(\underline{x})$ is quadratic function and $\underline{f}(\underline{x})$ and $\underline{g}(x)$ are linear functions.

e.g. $p(\underline{x}) = a + bx_1^2 + x_2^2$

linear

nonlinear

$x_1 x_2$ is also quadratic.

You can convert non linear system to linear by lineariz.

with inner loop.

⇒ Local and global optima. :-

interested in

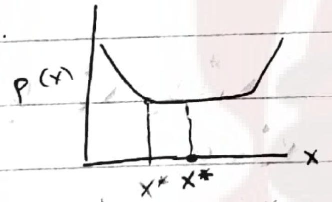
- In general we are interested in global optima.

Min $p(\underline{x})$ suppose that the optima occurs at $\underline{x} = \underline{x}^*$

Definition

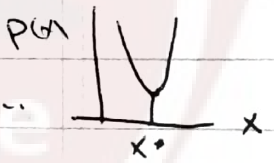
- x^* is local minimum if $p(\underline{x}) \geq p(\underline{x}^*)$ for all $\underline{x}$ in the neighborhood of x^*.

better than



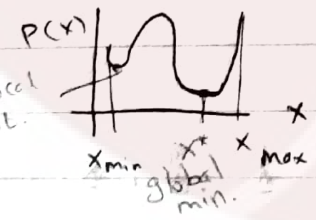
You have margin: may be this x^* is better.

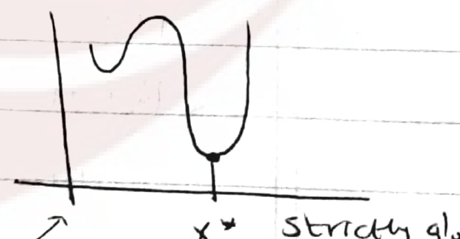
- x^* is strictly local min. if $p(\underline{x}) > p(\underline{x}^*)$ for all " "



- x^* is global min. if (iff) $p(\underline{x}) \geq p(\underline{x}^*)$ for all $\underline{x}$ in the whole Feasible region.

$x_{min} \leq x \leq x_{max}$





- x^* is strictly global iff $p(\underline{x}) > p(\underline{x}^*)$ — " "

- Feasible region

the region in which x^* satisfy all the constrains.

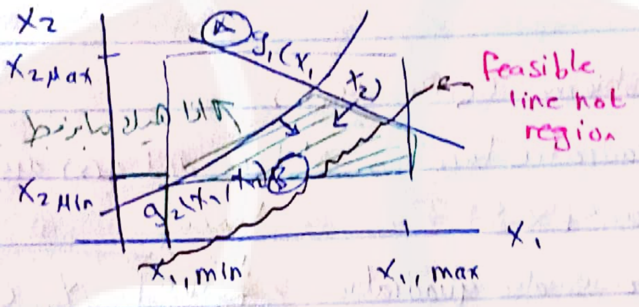
~~Feasible point.~~ لو اذا بدل تركيبة x_2 as fun (x_1) او اذا رح تشوف الكانفيشن ... دور الحو

min. $p(x_1, x_2) = \cdots$ subject to (s.t) Ⓐ $g_1(x_1, x_2) \geq 0$

Ⓑ $g_2(x_1, x_2) \geq 0$

Ⓑ $x_{1min} \leq x_1 \leq x_{1max}$

$x_{2min} \leq x_2 \leq x_{2max}$



Ⓒ $f(x_1, x_2) = 0$

$x_2 - 2x_1 + 3 = 0$

$x_2 = 2x_1 - 3$

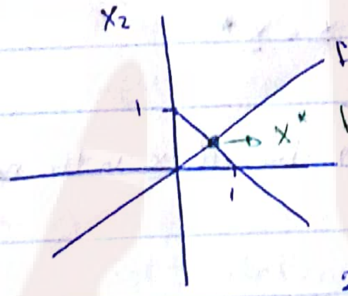
domain بدل تحد, جوه او ركن الرح

$g_1(\underline{x}) = x_2 + e^{-0.1x_1} \geq 0$

Example minimize $p(x_1, x_2) = x_1^2 + x_2^2$ s.t. $x_1 - x_2 = 0$ $f_1(x_1, x_2)$

constrained problem, there is not opt. $x_2 + x_1 - 1 = 0$ $f_2(x_1, x_2)$

variable must be bigger



feasible point

no ways for optimization.

x_1^*

$2x_1 - 1 = 0$ $x_1^* = \frac{1}{2}$ $x_2^* = \frac{1}{2}$

if no constrains $\to$ $x_1^* = x_2^* = 0$.

* # variables - # dependent equality constrains.

No opportunity for opt.

In optimization

$\to$ We need derivative information first, second

$P(\underline{x})$

✓ Notation. in this course

- representation of column vector = $\underline{f}$, $\underline{g}$, $\underline{x}$ بنشيل تحت خط

Rank $\underline{f} = [\quad]_{m_1 \times 1}$

$\underline{g} = [\quad]_{m_2 \times 1}$

$\underline{x} = [\,]_{n \times 1}$

بنستخدم ال optim derivative of

- representation of matrix (n x n) matrix (square matrix)

$\underline{\underline{A}}$, $\underline{\underline{H}}$

- ~~representation of scalar function~~

Derivatives of scalar function ($P(\underline{x})$) with respect to vector of variables ($\underline{x}$)

↳ 1st derivative (gradient)

مشتقة (اي زي بالنسبة لمسافة / لوقت ← تعطي rate

دل ادرج P $\nabla_{\underline{x}} P(\underline{x}) = \left[\frac{\partial P}{\partial x_1} , \frac{\partial P}{\partial x_2} \dots \frac{\partial P}{\partial x_n} \right]$

As row. if you take transpose it become column.

$\nabla P(\underline{x})\Big|_{x^*} = \underline{0}^{T}$ to make consistency.

Monday Optimization - 4/1/2019

- 1st derivative of scalar function p(x)

(gradient)

$$\nabla_{\underline{x}} p(\underline{x}) = \left[\frac{\partial P}{\partial x_1}, \frac{\partial P}{\partial x_2}, \cdots, \frac{\partial P}{\partial x_n} \right]$$ "raw matrix"

- 2nd derivative of scalar function $p(\underline{x})$

same

$$\nabla^2_{\underline{xx}} p(\underline{x}) = \begin{bmatrix} \frac{\partial^2 P}{\partial x_1^2} & \frac{\partial^2 P}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 P}{\partial x_1 \partial x_n} \\ \frac{\partial^2 P}{\partial x_2 \partial x_1} & \frac{\partial^2 P}{\partial x_2^2} & \cdots & \frac{\partial^2 P}{\partial x_2 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 P}{\partial x_n \partial x_1} & \frac{\partial^2 P}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 P}{\partial x_n^2} \end{bmatrix}$$

square matrix n x n

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

"Hessian matrix"

[≡] $\underline{\underline{H}}(\underline{x})$

always 2 lines.

$\underline{\underline{H}}(\underline{x})$ is a symmetric matrix : $\underline{\underline{H}}(\underline{x}) = \underline{\underline{H}}^T(\underline{x})$

it is transpose

الصف نفس العمود

imp in multivariables

- 1st derivative of vector function $\underline{F}(\underline{x})$ w.r.t vector of variables.

$$\underline{F}(\underline{x}) = \begin{bmatrix} F_1(x_1, x_2, \cdots, x_n) \\ F_2(x_1, x_2, \cdots, x_n) \\ \vdots \\ F_m(x_1, x_2, \cdots, x_n) \end{bmatrix}$$

$$\nabla_{\underline{x}} \underline{F}(\underline{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}_{m \times n}$$

n x m matrix

Jacobian matrix $\underline{J}(\underline{x})$

Scanned by CamScanner

Examples

Formulation of the optimization problem (converting the word statement of the problem to mathematical equations/model)

① $\boxed{1.5}$ optimal scheduling (common in plant design)

How many days per year (365 days) should each plant processes each product in order maximize the annual profit?

the table contain constants

→ Define the variables

t_{A1} : # days per year plant A operates product 1 ; [day/year]

t_{A2} : " " 2 "

t_{B1} : " B " 1 "

t_{B2} : " B " 2 "

→ Define the constant coefficients (investigation بعمل تحليل سوقي)

M_{A1} M_{A2} M_{B1} M_{B2} S_{A1} S_{A2} S_{B1} S_{B2}

ع مصنع
ع منتوج
ع تشغيل

→ Objective Function (Define)

$S_{A1} * M_{A1} \; [=] \; \frac{\$}{lbm} \times \frac{lbm}{Year} \; ; \; \frac{day}{year}$?

other factors → constrains

maximize profit $\equiv p(\overset{x_1}{t_{A1}}, \overset{x_2}{t_{A2}}, \overset{x_3}{\cancel{t_{A1}}} t_{B1}, \overset{x_4}{t_{B2}})$.

$= S_{A1} M_{A1} t_{A1} + S_{A2} M_{A2} t_{A2} + S_{B1} M_{B1} t_{B1} + S_{B2} M_{B2} t_{B2}$

objective function is linear

بس تغير كل
ضغة تبدل.
but doesn't make sense

→ Define constrains.

① Variable bounds ?

can be cancelled • $0 \le t_{A1} \le 365$ • $0 \le t_{A2} \le 365$ • $0 \le t_{B1} \le 365$ • $0 \le t_{B2} \le 365$

it is preferred to minimize constrains.

② equality constrains

$t_{A1} + t_{A2} = 365$

$\cancel{t_A}\; t_{B1} + t_{B2} = 365$

- constrains come explicitly or implicitly.
- Constrains from further analysis? Yes.

market demand $M_{A1} t_{A1} + M_{B1} t_{B1} \le L_1$ ← constant of demand. = constant

$M_{B2} t_{B2} + M_{A2} t_{A2} \le L_2$ = constant (another constant)

* This problem is linear optimization problem

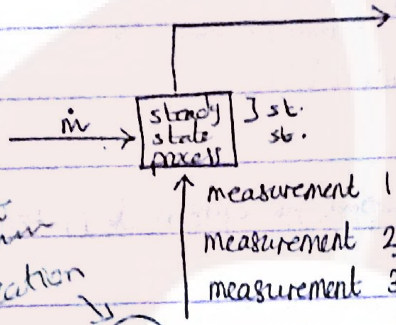
- it can be non linear from profit fluctuation (S becomes variable) and thus 8 variables of materials proceeds

اذا كان تغير
بسيط

↳ You can make sensitivity analysis instead of putting variables.

② 1.6

measurement 1   92.4 kg/h
measurement 2   94.3 kg/h
measurement 3   93.8 kg/h



measurement 1 : 11.1 kg/h
measurement 2 : 10.8 kg/h
measurement 3 : 11.4 kg/h

You can take average value but it is not optimization.

or < worst / optimum — optimization

Find the (best) value of $\dot{m}$

معناها بالكلاسالب الموجب

→ Define objective Function : sum of (square) of errors between some of inlet and outlet of the mass flow rate

Non linear.

. min of error

$$\text{Min. } p(\dot{m}) = (\dot{m} + 11.1 - 92.4)^2 + (\dot{m} + 10.8 - 94.3)^2 + (\dot{m} + 11.4 - 93.8)^2$$

$$\cancel{\partial P} \quad \frac{dP}{d\dot{m}} = 0 \qquad \text{find } \dot{m}$$

$$\dot{m}^* = 82.4 \text{ kg/h}.$$

Solve in avg. → the same value. / نفس الاشي حل معادلة ~~...~~ خطية غير خطية.

if accurate → error = 0

opti. ⌐ Partial differentiation حل معادلة ⇒ multivariable solving

Go to chapter 1 & solve.

Scanned by CamScanner

impirical
< model
based

Wednesday Optimization 6/11/2019

CHAPTER #2 : Modeling for optimization

→ Mathmatical model : (MA)

set of equations (s) that represent system of interest.

→ Solution of the (MM) :

the responsed the system at certain conditions.

→ why do we need M.M?

- enhance the understanding of the problem → Parametric study

  E.x does Temp affect the conversion in batch reactor.

- enhance the prediction accuracy

  E.x what is the T and P that give conversion 80%.

  what is the power of the pump to transport liquid
  from one location to another.]

  mass balance & energy balance & thermodynamic

  ① $\dot{m} = \rho Q$

  ② mechanical energy balance

  ③ darcy equation

  ④

→ How usefull is MM?

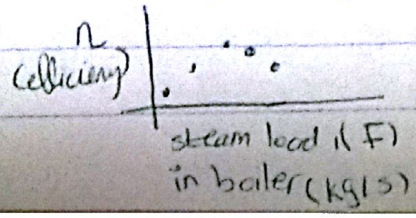
- evaluate different scenarion without having a process :→ (before) plant design
- هل انا بحاجة للتركيب او لا evaluate peformence before implementation (تركيب ، تشغيل) → Process control
- reduce experimentation : saving money and time.

→ Types of MM :-

not all models
models are not perfect

Ⓐ Empirical (experimental) models :

- developed from experiments
- engineering knowledge of the system is not necessary required



$\eta = a + bF + cF^2$

Find a, b, c by least square method as an optimization technique to get the best values of a, b, c.

Scanned by CamScanner

F h

1

2

3

4

5

مهمة

Ⓑ Fundemental models ✓ — conservation

(Fnn)

- developed From physical / chemical / biological concept. (rxn) (cell population balance)
  - e.g. mass balance, energy balance, thermodynamics (EOS, ...), cell population balance, --, etc.
- require knowledge of the system

Fundenental models

موزع بتوزيع space — Distributed models
Focus on the variation in space.

Lumped models. ~~time~~ no change in space.
Focuses on the variation in time

change (unsteady) Dynamic — ~~Static~~ steady.

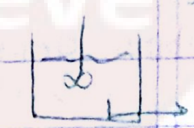
more than 2 variables (diff.)
independent variables: time, space
at least 1 dimension in space

ODE PDE — without time (no change in time).
e.g. PF reactor

can't by ordinary PDE

Dynamic: differential O.D.E

Steady: algebric
e.g. CSTR design equation



بنفترض انه بال اساس
good mixing 9c
to check imperically
لما تمر الأسكتشة .

building

~~→~~ modeling ~~development~~

optimization ليس ضروريا على

Ⓐ Problem definition phase

- define goal(s) (objective(s))
- state assumption (safe) (as much as you can)
- define variables and parameters

Ⓑ Model development phase

$A+B \rightarrow C$ سريع

$X_A = 100\% \rightarrow t = \infty \rightarrow$ ازالي $V = \infty$

Not safe if r slow rxn reversible

- define the system or subsystem for modeling.
- write the mathmatical equation(s)
- analyze the degree of Freedom

Ⓒ Solve the mathmatical model

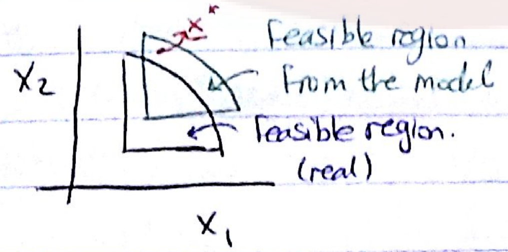
Imp. Ⓓ model validation phase التأكد من صحة النتائج
(make sense)

- Engineering knowledge
  - the results make sense?
  - Correct trend
- Compare with known results (either experimental or published data)

~~• Published data~~ بشغل ال model بعمل كل ظروفهم وبقارن على تجربة وات كثيرة بنشغل فنقارن

calibrate thermometer مع ميزان زئبقي ← - calibrate and validate model.

Rule No model is perfect — to 10% error it is acceptable. maybe the assumptions are not safe.



** Degree of freedom DoF

$\# DoF = n - m$

#variables ← #independent equations ← by rank

$\underline{\underline{A}} \rightarrow |A|$: Determinant

$m = rank(\underline{\underline{A}})$

E.X $\underline{\underline{A}}\,\underline{x} = \underline{b}$

$\underline{\underline{x}} = n \times 1$

$b = m \times 1$

$\underline{\underline{A}} = n \times m$

Rank : highest order of square matrix or submatrix of A with non-zero determinant ($|A|$)

ازا أي س Submatrix

محددة non zero بتبني ال rank.

0 = #DoF : unique solution
no oppurtinity for optimization.

#DoF > 0 : there is opportunity for optimization and the p(x) comes to get one unique solution.

#DoF < 0 : no solution.

Monday 11/2/2019

consider $\underline{F}(\underline{x}) = 0$

where $\underline{x}$ has an order of $n \times 1$

$DoF = n - Rank(\nabla \underline{F}(\underline{x}))$

Ex is there an opportunity to ~~maximize~~ minimize

$p(\underline{x}) = x_1^2 + x_2^2$

s.t : $2x_1 + 4x_2 + 6x_3 = 1 \rightarrow f_1 = 2x_1 + 4x_2 + 6x_3 - 1 = 0$

$x_1 + 3x_2 + 5x_3 = 0 \rightarrow f_2 = x_1 + 3x_2 + 5x_3 = 0$

$9x_1 + 10x_2 + 11x_3 = 7 \rightarrow f_3 = 9x_1 + 10x_2 + 11x_3 - 7 = 0$

$n = 3 \quad \nabla(\underline{F}(\underline{x})) = \begin{bmatrix} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 9 & 10 & 11 \end{bmatrix} = \underline{A}$ Rank = 3

$|\underline{A}| = 0 \Rightarrow$ They are dependent

اذا واحد طلعت ال determinate non zero وقف

$\underline{A_1} = \begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix} = 2 \neq 0$ Rank = 2 (order 8 2x2)

$DoF = 3 - 2 = 1 > 0$ ✓ optimization

⟹ Unconstrained single variable optimization problem :

Max or Min of $p(X)$

→ why unconstrained single variable problem?

① some real engineering problems are unconstrained.

Ex. • least square method

• Thermodynamics

② sometimes unconstrained single variable problem is sub problem in constrained family.

بتعويض المتغيرات بدلالة الآخر

③ every optimization algorithm includes "line search" in constrained problems

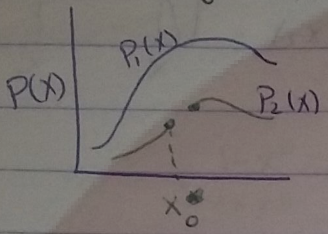
optimum distance in step moving

⇒ Mathmatical review

- continuity of $p(x)$

النهاية من اليمين تساوي من الشمال

$$\lim_{x \to x_0} p(x) = p(x_0)$$



in this course, we focus on continuous functions.

القيم الموجودة في السوق discrete

إذا كان $(P(D^+) = 3.21 \text{ (in)})$ لانستطيع أن نحصل عليه.

- Differentiability of $p(x)$

$$\lim_{\Delta x \to 0} \frac{p(x+\Delta x) - p(x)}{\Delta x} \text{ must exists.}$$

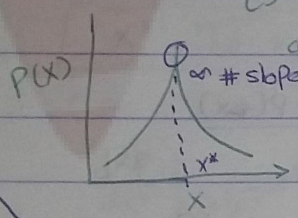
• كل اقتران differ. هو continuous لكن العكس غير صحيح دائما

at critical point (edges) differential does not exist

x^* is critical point.

طريقة الاشتقاق و مساواة بالصفر يعطي diverge

عشان هيك منأخذ طرق أخرى



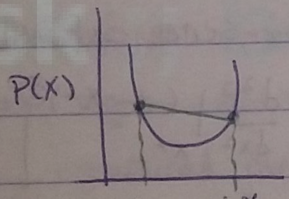
$\lim p(x+\Delta x)$ & $\lim p(x)$ must exist.

• every differentiable function is continuous, but the reverse is not always true

• if $p(x)$ is continuous but not differentiable, the optimization methods based on using derivatives will fail to find the optimum.

- Concave and convex ∪ ∩ (convexity of $p(x)$):-

ⓐ Convex Functions



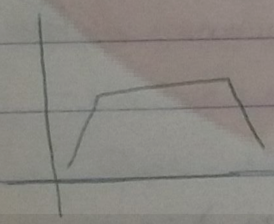
• it has minima.

• $p(x)$ is convex if $p(\lambda x_1 + (1-\lambda)x_2) \leq \lambda p(x_1) + (1-\lambda) p(x_2)$ on $[x_1, x_2]$ فترة

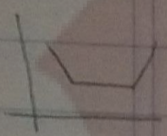
where λ is constant and has values $0 \leq \lambda \leq 1$

(take any value and substitute above)

قيمة الاقتران الأصلي أقل من قيمة ال linear

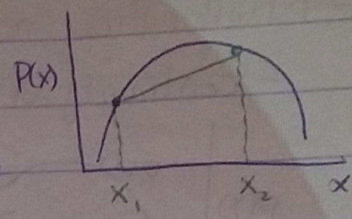


not strictly convex



convex but not strictly convex.

Ⓑ Concave function $\Rightarrow$ maxima.



- $P(x)$ is concaved on $[x_1, x_2]$ if

$$P(\lambda x_1 + (1-\lambda)x_2) \geq \lambda P(x_1) + (1-x)P(x_2)$$

- if you replace $\leq$ by $<$ it is called strictly convex.

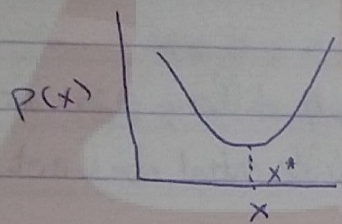
Previous ways (convexity) sometimes not practical

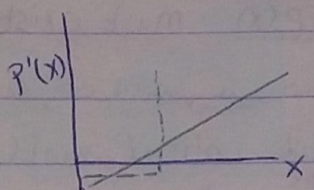
لكن مفكر للاستقرار
اذا الاقتران غير
differential
اصبح λ من
0 لـ 1

- optimality condition

if x^* is a minimum point then $p(x^*) < P(x^* + \Delta x)$ $\forall$ $\Delta x \neq 0$

تغيير الإشارة





slope of tangent $\leftarrow$ $\dfrac{dP}{dx} = \lim_{\Delta x \to 0} \dfrac{P(x+\Delta x) - P(\Delta x)}{\Delta x}$

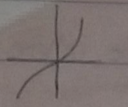
Δx around x^*, $\Delta x > 0$ $\quad dP/dx > 0$
$\qquad\qquad\qquad\;\;\Delta x < 0$ $\quad dP/dx < 0$ $\Big\}$ $\left.\dfrac{dP}{dx}\right|_{x^*} = 0$ "Necessary condition" is there optimum or not?

same for concave but different signs.

كل ال course ح نشتغل على .min مش .max ولو كان موجه لـ .min

"sufficient condition" :-

$\left.\frac{d^n P}{dx^n}\right|_{x^*} \neq 0$ If n is → odd → x^* is an inflection point
→ even → x^* is minimum if $\left.\frac{d^n P}{dx^n}\right|_{x^*} > 0$ vise versa



بسبب الأس حددت هنا الاختلاف
اما زوجي يعني السالب ، اذا فردي 8

$1 - x^4$ → maxima
$1 + x^4$ → minima

Example Min. $P(x) = 6 + 2x + \frac{3}{x^2}$ Optimiz. ⟨ time / accuracy

$\left.\frac{dP}{dx}\right|_{x^*} = 0$ $2 - \frac{6x}{x^4} = 2 - \frac{6}{x^3}$

$\boxed{x^* = \sqrt[3]{3} = 1.44224}$ exact soln. from analytical method

$\left.\frac{d^2P}{dx^2}\right|_{x^* = \sqrt[3]{3}} = \frac{18x^2}{x^6} = \frac{18}{x^4} = \frac{18}{\sqrt[3]{3}} > 0$ ✓ minima.

تطلع مشتقة على صفر أو ما الك دالة يدل على دالة.

dangerous if my function is not differentiable

Wednesday 13/2/2019

- Obtaining exact solution using analytical method is not always possible (when you can't find the derivative).

⇒ Other methods for solving single variable problem

Ⓐ method use objective function values

ما بدها مشتقات

- region elimination methods
- graphical method
- Polynomial interpolation methods

Ⓑ Methods use objective function values and its derivative values.

- Newtons method
- Quasi - Newton's method
- secant method

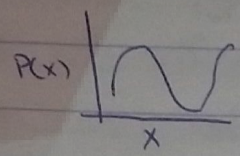
~~Wednesday~~

8 بعيدين في line search in multi

* Graphical method

المشكلة أنه ضايلنا local مو global عشان نشتغل كثير د excel

- Plot $P(x)$ on an interval of x and observe where is minima / maxima



Previous example

@ zero → a don't plot in this range

* Region elimination methods

نفس المشكلة بكن السحب أحل Physically احنا أمرق بدنا الحدود للتطبيق العملي

1. Find an interval which brackets the optima.
2. reduce the size of the interval
3. repeat 2 until converge to optima.

① Scanning and bracketing method

1. choose an initial guess value of the optima $k=0 \rightarrow x=0$
2. increase / decrease the value of x and evaluate $P(x_k)$

not uniform → $x_{k+1} = x_k + \delta * 2^k$; $k = 0, 1, \ldots$

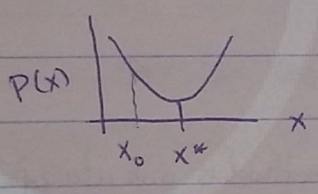
x (growth of Δx)

where δ is a constant parameter

$0 < \delta \leq 1 \rightarrow$ increasing (Δx +ve)

$-1 \leq \delta < 0 \rightarrow$ decreasing (Δx -ve)

make investigation on value

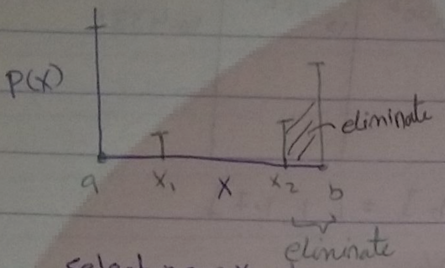


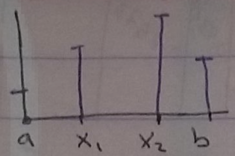
$k \quad x_k \quad P(x_k)$

one of them must be smallest

$P(a)$, $P(b)$, $P(x_1)$, $P(x_2)$ * $a < x_1 < x_2 < b$

how to distribute x_1 & x_2





~~x_2^* as new value~~

→ bad choice of interval

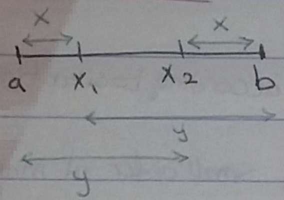
must choose point < a

select new x_2, b

جهة التقدم من اليمين أو الشمال

• How to select x_1 & x_2 ?

Use "Golden-section" ratio (r)



$L = a - b = x + y$

$\frac{y}{x+y} = \frac{x}{y} = r$ ← ratio $x < y$ gradual cutting

$\frac{y}{x+y} = r \rightarrow \boxed{y = r(x+y)}$ ⊛ $\frac{x}{y} = r \rightarrow \boxed{x = ry}$ substitute in ⊛

$y = r(ry + y)$ $y = r^2 y + ry$ Divide by y $1 = r^2 + r$

$r^2 + r - 1 = 0$

$r = \frac{1 \pm \sqrt{(1)^2 - (4)(-1)(1)}}{(2)(1)} = \frac{-1 \pm \sqrt{5}}{2} = \boxed{r = 0.618}$ ✓

$= -1.618$ & I want +ve distance

Now $\frac{y}{x+y} = 0.618 = \frac{y}{L}$ $\boxed{y = 0.618L}$

$$\boxed{\begin{aligned} x_1 &= b - y \\ x_2 &= a + y \end{aligned}}$$

• The interval is reduced by $(1 - 0.618) = \boxed{0.382}$ (You can prove it) make ~~regeneration~~

K	a	X_1	X_2	b	L	y	P(a)	$P(X_1)$	$P(X_2)$	P(b)
	1	1.2674	1.4326	1.7	0.7	0.4326	11	10.4024	10.3269	10.4381
	1.2674	1.4327	1.5347	1.7	0.4326	0.2673				
	1.2674	1.3695								

applying on previous example $[a, b] = [1, 1.7]$

when to stop ? ~~when (new value - old)/~~

$$\tilde{x}^* \cong \frac{x_1^* + x_2^*}{2}$$

• Convergence Criteria

غير مفضلة • $|\tilde{x}^*_{(k+1)} - \tilde{x}^*_{(k)}| < e_1$ STOP according to situation غير مفضلة .

OR

مفضلة • $\frac{|\tilde{x}^*_{(k+1)} - \tilde{x}_{(k)}|}{\tilde{x}_{(k)}} < e_2 = 0.001$ small order of magnitude STOP

OR

غير مفضلة • $|P(x_{(k+1)}) - P(x_k)| < e_3$ STOP

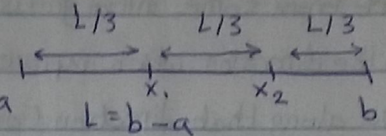
OR

مفضلة • $\frac{|P(x_{k+1}) - P(x_k)|}{P(x_k)} < e_4 = 0.001$ STOP

- can be found by calculator

* Two internal points with equal subintervals
→ reduction by 1/3 of the full interval
[a b] in which there is optima.

طريقة الفصل احسن



$L = b - a$

$x_1 = a + L/3$

$x_2 = a + 2L/3 = b - L/3$

Make error all abs $\frac{\text{abs(new - old)}}{\text{abs old}}$	between x^* $x^* = \frac{\cancel{a+b}}{2}$ $\frac{x_1 + x_2}{2}$

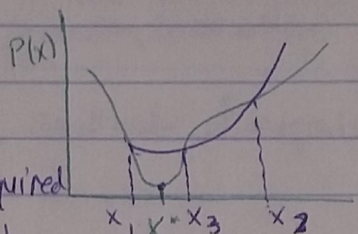
② Polynomial interpolation methods

أقل عدد iterations

$P(x) \cong Q_n(x) \quad n \geq 2$

IF $n = 2 \rightarrow Q_2(x)$: quadratic interpolation

$P(x) \cong Q_2(x) = a + bx + cx^2$ 3 points required



- choose 3 points (x_1, x_2, x_3) that bound the optima. $P(x_1) = Q_2(x_1) = a + bx_1 + cx_1^2 \rightarrow$ ①
- Convert you equation from dimentional to dimensionless to have values of a, b, c between 0 and 1 (choosing interval)

مثال if Temp from 20 to 200 make new variable = $\frac{-20+T}{200+T}$

$p(x_2) = Q_2(x_2) = a + bx_2 + cx_2^2 \rightarrow$ ②

$p(x_3) = Q_2(x_2) = a + bx_3 + cx_3^2 \rightarrow$ ③

Solve to obtain a, b, c

$P(x) \cong Q_2(x) \rightarrow P'(x) \cong Q_2(x) = b + 2c\hat{x}^* = 0$

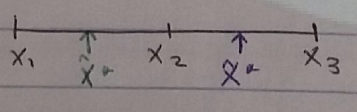
$\hat{x}_* = \frac{-b}{2c}$

You can verify it ! $\boxed{\hat{x}_* = \frac{1}{2}\left[\frac{(x_2^2 - x_3^2)P(x_1) + (x_3^2 - x_1^2)P(x_2) + (x_1^2 - x_2^2)P(x_3)}{(x_2 - x_3)P(x_1) + (x_3 - x_1)P(x_2) + (x_1 - x_2)P(x_3)}\right]}$

where $x_1 < x_2 < x_3$

only if Q(x) is quadratic $P(x) = Q(x)$

in the next iteration



$x_1 = x_1$, $x_2 = \hat{x}^*$, $x_3 = x_2$

$x_1 = x_2$, $x_2 = \hat{x}^*$, $x_3 = x_3$

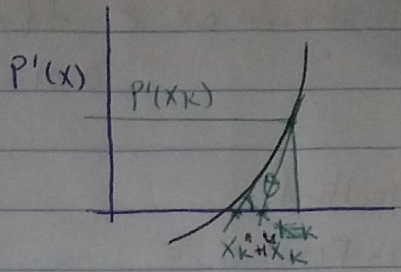
- Use previous methods that use objective function value ONLY for non-smooth p(x) since they have slow rate of convergence
- Faster rate of convergence is possible using the second family that use of values and their derivatives.

① Newton's method

Generally $F(x) = 0 = P'(x)$

Here $p''(x) = F(x) \cancel{\neq 0}$

$$F'(x) = \tan\theta = \frac{P'(X_k)}{X_k - X_{k+1}}$$



$$X_k - X_{k+1} = \frac{P'(x)}{F'(X_k)} =$$

$$\boxed{X_{k+1} = X_k - \frac{P'(X_k)}{P''(X_k)}}$$

OR derive it from Taylor series expansion.

k	X_k	$P(X_k)$	$P'(X_k)$	$P''(X_k)$	Error %
0	x_0	✓	✓	✓	–

To compare compare with 1.35 not 1

x_0 initial guess value

Taylor series

$$P(x) \cong P(X_k) + p'(X_k)(X - X_k) + \frac{p''(X_k)}{2}(X - X_k)^2$$ Trunkate others

$$P'(x) = 0 = 0 + p'(X_k) + \frac{2}{2} p''(X_k)(X - X_k)$$

$$\frac{-P'(X_k)}{P''(X_k)} = X - X_k$$

$$\boxed{X = X_k - \frac{P'(X_k)}{P''(X_k)}}$$

newton's method as approximation is based on quadratic Polynomial

may diverge if far initial guess from the correct value

② Quasi - Newton's method

Analytical derivatives are used, if you don't have the function or not smooth function ⇒ use quasi Newton's method.

$$X_{k+1} = X_k - \frac{P'(X_k)}{P''(X_k)}$$

if root is True ⇒ $X_{k+1} = X_k$

(درجة ثانية الدقة) central

- it uses Finite difference Formulas to approximate $p'(X_k)$ and $p''(X_k)$

$$p'(X_k) \cong \frac{P(X_k + \Delta X) - P(X_k - \Delta X)}{2\Delta X}$$ X_{k-1}

X_k

$$p''(X_k) = \frac{P(X_k + \Delta X) - 2P(X_k) + P(X_k - \Delta X)}{(\Delta X)^2}$$ X_{k+1}

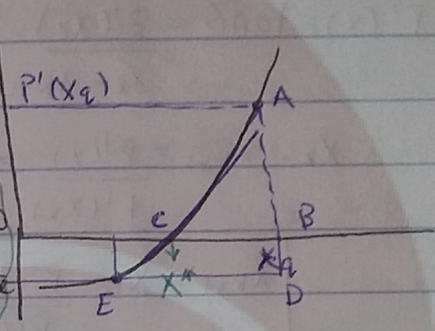
• How to decide Δx ? depending on order of magnitude of the numbers.
0.01 * order of magnitude

③ Secant method

$P'(x) = F(x) = 0$

it requires two points : x_p & x_q

that bracket the optima such that $P'(x_p)P'(x_q) < 0$ ⇓ intermediate value theorem



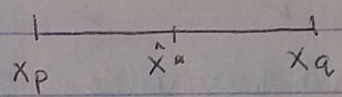
similar triangles : ABC and ADE

تقريب الحقيقة ←

$$\frac{x_q - \hat{x}^*}{x_q - x_p} = \frac{P'(x_q) - 0}{P'(x_q) - P'(x_p)}$$

$$\boxed{\hat{x}^* = x_q - \frac{P'(x_q)(x_q - x_p)}{P'(x_q) - P'(x_p)}}$$

ممكن نتخيلها للطرق السابقة



if $p'(\hat{x}^*) \cdot p'(x_q) < 0$

then $x_p = \hat{x}^*$ $x_q = x_q$

and vice versa.

* slower rate of convergence than Newton and quasi-Newton methods

Matlab command :

Fminbnd input a and b

bnd: interval

Unconstrained multi-variable optimization

Wednesday
20/1/2019

Min $p(\underline{x}) = \dots$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad x_1, x_2, \dots, x_n$$

** Optimality conditions

$\frac{dP}{dx}$ You can Divide function $\frac{dP_1}{dx} + \frac{dP_2}{dx} + \dots = 0$

* Necessary condition :

If $\underline{x}^*$ is a min. then $p(\underline{x}^*) \leq p(\underline{x}^* + \Delta\underline{x}) \ \forall \Delta\underline{x}$

using Taylor series expansion around x^* $\quad \Delta\underline{x} = \underline{x} - \underline{x}^* = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix}$

$$\underbrace{P(\underline{x})}_{\text{scalar}} = \underbrace{p(\underline{x}^*)}_{\text{scalar}} + \underbrace{\underbrace{\nabla_{\underline{x}} P(\underline{x})\big|_{x^*}}_{\text{order } 1\times n} \underbrace{(\underline{x} - \underline{x}^*)}_{\text{order } n\times 1}}_{\text{scalar}} +$$

You can't write $(\Delta x)^2$
But $\underset{n\times 1}{\Delta\underline{x}}$ $\underset{n\times 1}{\Delta\underline{x}}$ so
Take Transpose according to the order you want

not necessary Δx to be +ve or −ve

$$\underbrace{\frac{1}{2} \underset{1\times n}{\Delta\underline{x}^T} \underset{n\times n}{\nabla^2_{\underline{x}\,\underline{x}} p(\underline{x})} \underset{n\times 1}{\Delta\underline{x}^T}}_{\text{scalar}} + \dots$$

Heishian matrix

← must be +ve

Necessary condition $\nabla_{\underline{x}} p(\underline{x})\big|_{x^*} = \underline{0}^T$ must be as row

* sufficient condition for minima

$\Delta\underline{x}^T \nabla^2 p(\underline{x})\big|_{x^*} \Delta\underline{x} > 0$ OR the $\underline{\underline{H}}(x)\big|_{x^*}$ is positive definite

How?

- The geometry of $p(x)$ is characterized by $\underline{\underline{H}}(x)$ # values = # n

easier to check than ←

$\underline{\underline{H}}$	$\Delta\underline{x}^T \underline{\underline{H}} \Delta\underline{x}$	Eigen values of $\underline{\underline{H}}$	$p(\underline{x})$	$p(\underline{x}^*) - p(x^*)$	Geometry
① positive definite	> 0	all positive	strictly convex	increase	vally
② Positive semi-definite	≥ 0	some positive & some have zero	convex	possibly increase ممكن يضل ثابت	trough
③ negative definite	< 0	all negative	strictly concave	decrease	Hill
④ negative semi-definite	≤ 0	some −ve & some zeros	concave	possible decrease	Hill ridge
⑤ Indefinite	≥ 0 or ≤ 0 depending on $\Delta\underline{x}$	some +ve & some −ve	saddle point	increase or decrease or neither	Saddle point.

③ reflection نقطة في single variable

in reality يجب أن نذهب إلى المشتقة الثالثة لكن غير موجودة

وإذا كانت تساوي صفر نذهب الى المشتقة الرابعة.

- Example : classify the geometry of ~~p(x)~~ $p(x_1, x_2) = 2x_1^2 - 3x_1x_2 + 2x_2^2$ then find X^* by drawing contour of $p(x) = 1$ for x_1, x_2

ونجرب الي بيطلع على شكل عيون

$p(x) = 2$ for x_1, x_2

ويمكن أكثر من وحدة / القمة طويلة

$$\nabla P(\underline{x}) = \begin{bmatrix} \frac{\partial P}{\partial x_1} & \frac{\partial P}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x_1 - 3x_2 & 3x_1 + 4x_2 \end{bmatrix}$$

$$\nabla^2_{\underline{x}\underline{x}} P(x) = \underline{\underline{H}} = \begin{bmatrix} \frac{\partial^2 P}{\partial x_1^2} & \frac{\partial^2 P}{\partial x_1 \partial x_2} \\ \frac{\partial^2 P}{\partial x_1 \partial x_2} & \frac{\partial^2 P}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -3 & 4 \end{bmatrix}$$

if it is x^3 → matrix have variables, Find x^* and substitute in matrix

Eigen value

$$\begin{bmatrix} 4-\lambda & -3 \\ -3 & 4-\lambda \end{bmatrix} = 0 \qquad (4-\lambda)(4-\lambda) - 9 = 0 \qquad (4-\lambda)^2 - 9 = 0$$

$\lambda_1 = 7 \quad \lambda_2 = 1$ all +ve

$\underline{\underline{H}}$ is positive definite → $p(x)$ is strictly convex and has unique minima

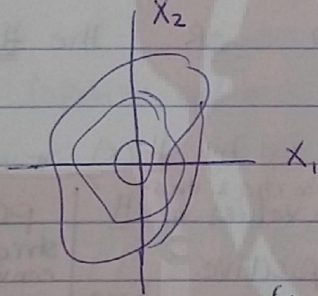
If you want to plot

$$\nabla p(x)\Big|_{\underline{x}^*} = 0^T \qquad \begin{bmatrix} 4x_1^* - 3x_2^* \\ -3x_1^* + 4x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$4x_1 - 3x_2 = 0 \qquad x_2^* = \frac{4}{3} x_1^* \qquad \underline{x}^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-4x_1^* + 4x_2^* = 0$$

PLOT



Example : minimize $p(\underline{x}) = \cancel{2} - x_1^3 \, e^{(x_2 - x_1^2 - 10(x_1 - x_2)^2)}$

$\nabla p(\underline{x}) = \underline{0}^T \qquad \frac{dP}{dx_1} = 0 \quad 0 = -3x_1^2 \, e^{(x_2 - x_1^2 - 10(x_1 - x_2)^2)} + -x_1^3$

$\frac{dP}{dx_2} = 0 \quad 0 = -3x_1^2 e$

high non linearity

solve the two highly non linear to get x^*

Remark The analytical is not always simple to deal with → alternative is the numircal methods

* Approach of numerical methods :

① select an initial guess point(s)

② Find the direction which improves the value of p(x)

③ move along that direction

④ repeat ② and ③ until converged.

→ The different numerical methods come from the way in which it uses the search direction.

* Approach of numerical methods:

1. select an initial guess point(s)
2. find the direction which improves the value of $p(x)$
3. move along that direction
4. repeat 2 and 3 until converged.

→ The different numerical methods come from the way in which it uses the search direction.

25/2

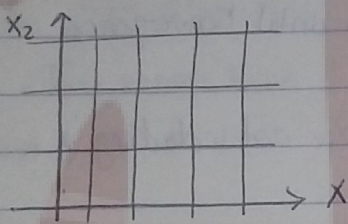
* Numerical methods for unconstrained multi-variable optimization function

1. Random search method

min. $p(\underline{x})$

x_1	x_2	x_3	... x_n	$p(\underline{x})$

mesh (grid)



100 point with 10 variables

↳ calculation of $p(x)$

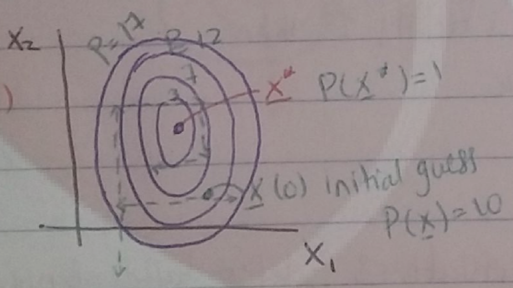
→ Huge calculation of $p(x)$ when n is large (~~pat~~ but it can be).

2. Univariate search method

→ search for the optimum in each direction independently

نثبت كل شيء ونغير واحد فقط

Min $p(\underline{x})$



بس تلاقي $p(x)$ تزيد وتنقص

ويروح المتغير الآخر هكذا

"orthoganity"

→ effective for $p(\underline{x})$ of the form $p(x) = \sum_{i=1}^{n} c_i x_i^2 \rightarrow \underline{x}^* = 0^T$

if $x_1, x_2 \rightarrow$ circles

$x_1, x_2, x_3 \rightarrow$ balls

→ less effective for $p(\underline{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_i x_j$ ellipse

~~constant + x_1 + $x_1 x_2$~~ $\quad Ax_1^2 + Bx_2^2 + C(x_1 x_2)$

Most of softwares

② Simplex method (in matlab)

- simplex : geometric figure for by n+1 points in the n-dimensional space   No orthogonality

  if $x_1, x_2 \rightarrow 3$ points $\rightarrow$ triangle

  $\cdot P_2$

  $P_3 \cdot$ I know were is the worst point.

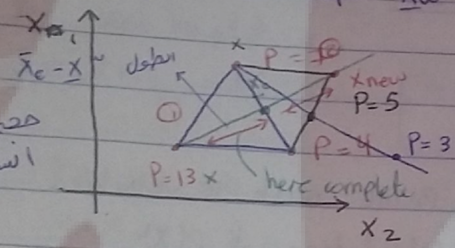
  $\cdot P_1$ (initial guess)

Procedure

1. start with a simplex of n+1 vertices, where $\dim(\underline{x}) = n$
2. evaluate $P(\underline{x})$ at each vertex and find the worst point ($\underline{x}_w$) (which has the highest $p(\underline{x})$)
3. reflect the $\underline{x}_w$ through the centroid of the simplex formed by all the points except $\underline{x}_w$.

   centroid: middle between the other two points

من الغلط بالتجريب



حجم المثلث هيصغر كلما استقرينا.

Reflection — complete / expansion / contraction   أنت بالتجريب

And so on until convergence of x_{new} with x_{old}

4. Form the new simplex by rejecting the $\underline{x}_w$ and including $\underline{x}_{new}$
5. repeat 2-4 untill converged.

$$\boxed{\underline{x}_c = \frac{\sum_{i=1}^{n+1} \underline{x}_i - \underline{x}_w}{n}}$$

$$\boxed{\underline{x}_{new} = \underline{x}_c + \alpha(\underline{x}_c - \underline{x}_w)}$$

if α (depends on $p(x)$):
- $= 1$ complete reflection
- > 1 reflection and expansion
- < 1 reflection and contraction

- simplex method has slow rate of convergence (only finding 1 of f(p(x)) each iteration)
  - You make investigation on α, firstly 1, then iterations $\rightarrow$ better α value

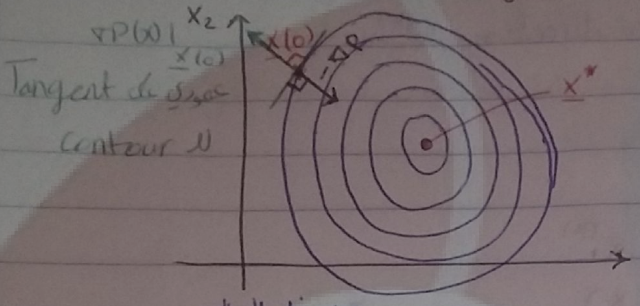
In Matlab fminsearch

Until know they don't depends on derivatives., let's take derivatives

** Gradient based method (it is raw vector)

$\nabla p(\underline{x})$

→ The direction of $\nabla p(\underline{x})$ is orthogonal (perpendicular) to the $p(\underline{x})$ contours at any point.



ما دخل بالسابق

→ a descent direction $\underline{s}$ which direct to minima must satisfy

dot product يدل على الاتجاه $\quad \nabla p(\underline{x}) \cdot \underline{s} \leq 0 \qquad$ if $\underline{s} = -\nabla p(\underline{x})$

$\nabla p(x) \; s \; \cos\theta$

هنا $180°$ المحصورة بين s و ∇p $\qquad 90 < \theta \leq 180$

اذا اقل من $90°$ سيء

• This means that if $\theta = 90°$: no improvement in optima.

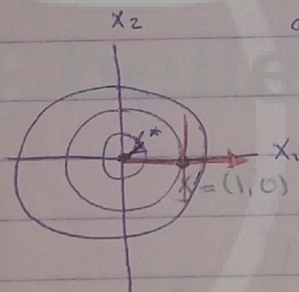
if $0 \leq \theta \leq 90°$: no improvement

if $90 < \theta \leq 180°$: improvement of optima (minima)

Example $\quad$ Min $p(\underline{x}) = x_1^2 + x_2^2$

$x^* = [0 \;\; 0]^T \qquad \nabla P(\underline{x}) = [2x_1 \;\; 2x_2]^T$

at $\underline{x}^{(0)} = [1, 0]^T$: $\nabla P = [2 \;\; 0]$ $\quad$ خارج مهم 0 0



even at $[1 \;\; 1]^T$, gradient is ⊥ tangent

set $k = 0$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \underbrace{\lambda_k}_{\text{مقدار}} \underbrace{\underline{s}^{(k)}}_{\text{اتجاه}}$$

(need optimum value)

↓

line search

وبعدين منغير اتجاه اذا تطلب

Wednesday
27/2/2019

For minimization (maximization → ascent)

** steepest descent method :

→ initial guess point, $\underline{X}^{(0)}$ of $\underline{X}^{*}$ for minimization.

→ The search direction : $\underline{S}^{(k)} = - \nabla_{\underline{X}^{(k)}} P(\underline{X})$

→ Rule : $\underline{X}^{(k+1)} = \underline{X}^{(k)} + \lambda_k \underline{S}^{(k)}$ $k = 0, 1, 2, \dots$

λ_k : optimum step size along the direction $\underline{S}^{(k)}$

Algorithm

1. set $k = 0$
2. select initial guess point $\underline{X}^{(k)}$
3. determine $\underline{S}^{(k)} = -\nabla_{\underline{X}_k} P(\underline{X})$
4. perform line search to obtain ~~λk~~ X_k

على اتجاه واحد
unidirection
↓
single variable optimization

$\underline{X}^{(k+1)} = \underline{X}^{(k)} + \lambda_k \underline{S}^{(k)}$ only λ_k is unknown

$P(\underline{X}) = \dots$

$P(\underline{X}^{(k+1)}) = G(\lambda_k)$

$\dfrac{dP(\underline{X}^{(k+1)})}{d\lambda_k} = 0$ → Find λ_k^*

5. evaluate $\underline{X}^{(k+1)} = \underline{X}^{(k)} + \lambda_k \underline{S}^{(k)}$

Convergence criteria on P or X

6. If $\sum_{i=1}^{n} |x_i^{(k+1)} - x_i^{(k)}| / |x_i^{(k)}| < \epsilon$ stop

7. $k = k+1$
8. go to step 3

Example

$\min P(\underline{X}) = x_1^2 + x_2^2$

$k=0$

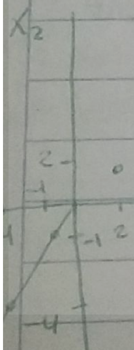
$\underline{X}^{(0)} = [2 \quad 2]^T$

$\nabla P(\underline{X}) = [2x_1 \quad 2x_2]^T$

$\nabla_{\underline{X}(0)} P(\underline{X}) = [4 \quad 4]^T$

$\underline{S}^{(0)} = -\nabla_{X(0)} P(\underline{X}) = [-4 \quad -4]^T$ You can take $[-1 \quad -1]^T$ because it is a vector

if you multiply or divide a vector by scalar



الافضل ان لا تغير الاشارة، عشان نخلي λ موجبة، ما بنمشي عكس الاتجاه

$$\underline{X}^{(1)} = \underline{X}^{(0)} + \lambda_0 \underline{S}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \lambda_0 \begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 - 4\lambda_0 \\ 2 - 4\lambda_0 \end{bmatrix}$$

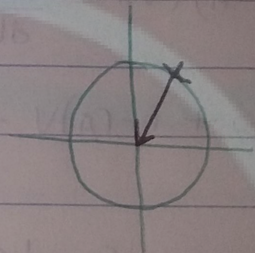
Line search : $P(\underline{X}^{(1)}) = P(x_1^{(1)}, x_2^{(1)}) =$

$P(2 - 4\lambda_0, 2 - 4\lambda_0)$

$$P(\underline{x}_1^{(0)}) = (2-4\lambda_0)^2 + (2-4\lambda_0)^2 = 2(2-4\lambda_0)^2$$

$$\frac{dP}{d\lambda_0} = 0 \rightarrow 4 * -4(2-4\lambda_0) = 0 \qquad 2-4\lambda_0 = 0 \qquad \boxed{\lambda_0 = \frac{1}{2}}$$

substitute in $\underline{x}^{(1)} = \begin{bmatrix} 2-4(0.5) \\ 2-4(0.5) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



المفروض انه ما يوقف هوة اذا بيقدر نحل analytical

ما نستخدم هاي الطريقة

Quiz

$A \rightarrow B$ $\quad\quad I - O + G - C = A$

$B \rightarrow C$

(A) $-(-r_A)(V) = \frac{dN_A}{dt}$ $\quad\quad -(-r_A) = \frac{dC_A}{dt}$

$0 - 0 + (-r_A)V - (r_c)V = \frac{dN_B}{dt}$

$-r_A - r_c = \frac{dC_B}{dt}$ $\quad\quad r_A = k_1 C_A$

non homogeneous D.E

كل الي تحتهم خط هم vector (عمود) ماعدا ∇

4/3/2019

- $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$ $\qquad$ λ_k نقدرها (1D)

- steepest method $\quad \underline{s}^{(k)} = -\nabla_{\underline{x}^{(k)}} P(\underline{x})$

- line search to get optimum λ_k : $\left.\dfrac{dP(\underline{x})}{d\lambda_k}\right|_{\underline{x}^{(k+1)}} = 0$

$$\left.\frac{dP(\underline{x})}{d\lambda_k}\right|_{\underline{x}^{(k+1)}} \overset{\text{chain rule}}{=} \left.\frac{dP(\underline{x})}{d\underline{x}}\right|_{\underline{x}^{(k+1)}} \left.\frac{d\underline{x}}{d\lambda_k}\right|_{\underline{x}^{(k+1)}} = 0 \qquad \underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$$

$\frac{d\underline{x}}{d\lambda_k} = \underline{s}^{(k)}$ order $n \times 1$

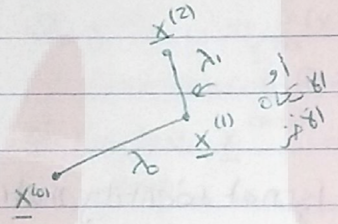
$\cancel{\partial P(\underline{x})}$ $\nabla_{\underline{x}^{(k+1)}} P(\underline{x})$ order $1 \times n$

$= -\underline{s}^{(k+1)}$

$= -\underline{s}^{(k+1)} \underline{s}^{(k)}$ equal zero when perpendicular

(vector x vector) → ضرب نقطي مع اتجاه $\qquad$ الزاوية المحصورة بينهم $\cos\theta$

So... $\underline{s}^{(k+1)} \perp \underline{s}^{(k)}$



* The orthogonality of the successive [$\underline{s}^{(k)}$ and $\underline{s}^{(k+1)}$] search directions leads to an inefficient zigzaging behaviour which slow down the rate of convergence.

** Conjugate search method

→ initial guess point $\underline{x}^{(0)}$

→ initial search direction $\underline{s}^{(0)}$ from physics or according to $-\nabla$

Definition $\qquad$ square matrix: مربعة ← w.r.t $\underline{\underline{H}}(\underline{x}^{(k)})$

Two directions $\underline{s}^{(k)}$ and $\underline{s}^{(k+1)}$ are said to be conjugated if

$$\underbrace{\underline{s}^{(k)T}}_{1\times n} \underbrace{\underline{\underline{H}}(\underline{x}^{(k)})}_{n\times n} \underbrace{\underline{s}^{(k+1)}}_{n\times 1} = \underbrace{0}_{1\times 1} \rightarrow \text{scalar}$$

Special case $\quad \underline{\underline{H}} = \underline{\underline{I}}$ identity matrix → $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

ضرب I في matrix نفس matrix $\qquad$ when ?

$\underline{s}^{(k)T} \underline{\underline{I}} \underline{s}^{(k+1)} = 0$ $\qquad$ Quadratic function $P(\underline{x}) = ax_1^2 + bx_2^2 + cx_3^2 + \dots$

$\underline{s}^{(k+1)} \underline{s}^{(k)} = 0$ They are perpendicular

I don't want that !

Scanned by CamScanner

Example $\quad \text{Min } p(\underline{x}) = 2x_1^2 + x_2^2 - 3 \qquad \underline{x}^{(0)} = [1 \;\; 1]^T \qquad \underline{s}^{(0)} = \underbrace{[-4 \;\; -2]}_{-\nabla_{\underline{x}^{(0)}} P(\underline{x})}$

K = 0

$\underline{x}^{(1)} = \underline{x}^{(0)} + \lambda_0 \underline{s}^{(0)}$

$\underline{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_0 \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1-4\lambda_0 \\ 1-2\lambda_0 \end{bmatrix} = \begin{bmatrix} -0.1111 \\ -0.4444 \end{bmatrix}$

line search to get λ_0

$\left.\dfrac{dP(\underline{x})}{d\lambda_0}\right|_{\underline{x}^{(0)}} =$

$p(x)\Big|_{\underline{x}^{(1)}} = 2(1-4\lambda_0)^2 + (1-2\lambda_0)^2 - 3$

$\left.\dfrac{dp(x)}{d\lambda_0}\right|_{x^{(1)}} = 0 \rightarrow (2)(2)(-4)(1-4\lambda_0) + (2)(-2)(1-2\lambda_0) = 0$

$\boxed{\lambda_0 = 0.27778}$

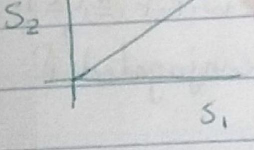
K = 1 $\quad \underline{x}^{(2)} = \underline{x}^{(1)} + \lambda_1 \underline{s}^{(1)}$

$\underline{s}^{(0)T} \, \underline{\underline{H}}(\underline{x}^{(0)}) \, \underline{s}^{(1)} = 0$

$\underline{\underline{H}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$ it is not identity matrix

الصف x العمود

$[-4 \;\; -2] \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} s_1^{(1)} \\ s_2^{(1)} \end{bmatrix} = 0$



$[-16 \;\; -4] \begin{bmatrix} s_1^{(1)} \\ s_2^{(1)} \end{bmatrix} = 0 \qquad -16 s_1^{(1)} - 4 s_2^{(1)} = 0$

عادي , معادلة بمجهولين

let $s_1^{(1)} = 1 \rightarrow s_2^{(1)} = -4$

ضع أي قيمة ما عدا الصفر

$\rightarrow \underline{s}^{(1)} = [-1 \;\; -4]^T \qquad \underline{x}^{(1)} = [-0.1111 \quad 0.4444]^T$

K = 2

$\underline{x}^{(2)} = \begin{bmatrix} -0.1111 \\ 0.4444 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ -4 \end{bmatrix} = \begin{bmatrix} -0.1111 + \lambda_1 \\ 0.4444 - 4\lambda_1 \end{bmatrix}$

Scanned by CamScanner

orthogonal when $\underline{\underline{H}}(\underline{X})$ is an identity matrix.

$$\frac{dP(\underline{x}^{(2)})}{d\lambda_1} = 0 \rightarrow 4(-0.1111 + \lambda_1) + (2)(-4)(0.4444 - 4\lambda_1) = 0$$

$$\boxed{\lambda_1 = 0.1111}$$

$$\underline{x}^{(2)} = \begin{bmatrix} -0.1111 + 0.1111 \\ -0.4444 - 4 \times 0.1111 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ as expected.

You can check if the matrix is Positive delinite or not

You can make new $g(x) = p(x) - 3$

gradient + Conjugate ** Conjugate gradient method (very fast)

Algorithm

① $K = 0$

② assume initial guess Point $\underline{x}^{(K)}$ ($\underline{x}^{(0)}$)

③ $\underline{S}^{(K)} = -\nabla_{\underline{x}^{(K)}} P(\underline{X})^T$

④ Line search $\longrightarrow \frac{dP(\underline{x}^{(K+1)})}{d\lambda_K} = 0$ to get λ_K

⑤ Compute $\underline{x}^{(K+1)} = \underline{x}^{(K)} + \lambda_K \underline{S}^{(K)}$

ما يميز الطريقة → ⑥ $\underset{n\times 1}{\underline{S}^{(K+1)}} = -\nabla_{\underline{x}^{(K+1)}} P(\underline{X})^T + \frac{\nabla_{\underline{x}^{(K+1)}} P(\underline{X}) \; \nabla_{\underline{x}^{(K+1)}} P(\underline{X})^T}{\nabla_{\underline{x}^{(K)}} P(\underline{X}) \; \nabla_{\underline{x}^{(K)}} P(\underline{X})^T} \underline{S}^{(K)}$ related to λ_K

excersise: solve the previous example by this way

⑦ IF $\sum_{i=1}^{N} \frac{|x_i^{K+1} - x_i^{K}|}{|x_i^{K}|} < e$ stop

⑧ $K = K + 1$

⑨ go to ④

if n

Find optimum (max / min للسؤال نقطة) ↓

Eigen values $\underline{\underline{H}}(\underline{X})$ نحسب

$-\lambda$ from diagonal element → Then determinate & equal it with zero

ممكن نطلع قيمتين لـ λ أو multiroots

$\nabla_{\underline{x}^{(k)}} p(\underline{x}) = \lambda_k \underline{s}^{(k)}$ Analog between $p(\underline{x}) = p(\underline{x}^{(k)}) + \nabla_{\underline{x}^{(k)}} p(\underline{x})(\underline{x} - \underline{x}^{(k)})$

$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$

→ Line search using quadratic approximation

→ $\underline{x}^{(k)}$

→ $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$

approximate $p(x)$ using Taylor series expansion around $\underline{x}^{(k)}$ with quadratic degree :

Taylor series

$$p(\underline{x}) \cong p(\underline{x}^{(k)}) + \nabla_{\underline{x}^{(k)}} p(\underline{x})(\underline{x} - \underline{x}^{(k)}) + \frac{1}{2}(\underline{x} - \underline{x}^{(k)})^T \underline{\underline{H}}(\underline{x})\big|_{\underline{x}^{(k)}} (\underline{x} - \underline{x}^{(k)}) + \dots$$

@ $\underline{x} = \underline{x}^{(k+1)}$ and using $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$

$$p(\underline{x}^{(k+1)}) = p(\underline{x}^{(k)}) + \nabla_{\underline{x}^{(k)}} p(\underline{x}) \lambda_k \underline{s}^{(k)} + \frac{1}{2} \lambda_k \underline{s}^{(k)T} \underline{\underline{H}}(\underline{x})\big|_{\underline{x}^{(k)}} \lambda_k \underline{s}^{(k)}$$

$$\frac{d\, p(\underline{x}^{(k+1)})}{d\lambda_k} = 0 \rightarrow 0 + \nabla_{\underline{x}^{(k)}} p(\underline{x}) \underline{s}^{(k)} + \lambda_k \underline{s}^{(k)T} \underline{\underline{H}}(\underline{x})\big|_{\underline{x}^{(k)}} \underline{s}^{(k)}$$

$$\boxed{\lambda_k = \frac{-\nabla_{\underline{x}^{(k)}} p(\underline{x})\, \underline{s}^{(k)}}{\underline{s}^{(k)T} \underline{\underline{H}}(\underline{x})\big|_{\underline{x}^{(k)}} \underline{s}^{(k)}}}$$

Imp:
gradient for scalar → vector
gradient for vector → matrix
$\nabla_{\underline{x}} \underline{x}$ → identity matrix $\underline{\underline{I}}$

6/3/2019

faster than steepest, may diverge

** Newton's method

→ initial guess point $\underline{x}^{(0)}$

→ quadratic expansion of $p(\underline{x})$ around $\underline{x}^{(k)}$ using TS function.

$$p(\underline{x}) \cong \underbrace{p(\underline{x}^{(k)})}_{\text{صفر}} + \nabla_{\underline{x}^{(k)}} p(\underline{x}) \underbrace{(\underline{x} - \underline{x}^{(k)})}_{\nabla_{\underline{x}^{(k)}} p(\underline{x})} + \frac{1}{2}(\underline{x} - \underline{x}^{(k)})^T \underline{\underline{H}}(\underline{x}^{(k)})(\underline{x} - \underline{x}^{(k)})$$

اشتق

* necessary condition $\nabla_{\underline{x}^*} p(\underline{x}) = \underline{0}^T$

ستحول ال numerical

$\underline{\underline{I}}$ * vector = vector

term وزعنا على كل ← $\nabla p(\underline{x}) = \underline{0}^T = \nabla_{\underline{x}^{(k)}} p(\underline{x}) + (\underline{x} - \underline{x}^{(k)})^T \underline{\underline{H}}(\underline{x}^{(k)})$

answer must be row, you must be consistent

@ $\underline{x} = \underline{x}^{(k+1)}$: $\underline{0}^T = \nabla_{\underline{x}^{(k)}} p(\underline{x}) + (\underline{x}^{(k+1)} - \underline{x}^{(k)})^T \underline{\underline{H}}(\underline{x}^{(k)})$

* matrix inverse doesn't change matrix order.

$$-\nabla_{\underline{x}^{(k)}} p(\underline{x}) = (\underline{x}^{(k+1)} - \underline{x}^{(k)})^T \underline{\underline{H}}(\underline{x}^{(k)})$$

Solve examples on Ch. ~~6~~

$$(\underline{x}^{(k+1)} - \underline{x}^{(k)})^T = -\nabla_{\underline{x}^{(k)}} P(\underline{x})\, \underline{\underline{H}}^{-1}$$

$$(\underline{x}^{(k+1)} - \underline{x}^{(k)}) = -\underline{\underline{H}}^{-1}\, \nabla_{\underline{x}^{(k)}} P(\underline{x}^{(k)})^T$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} - \underline{\underline{H}}^{-1}(\underline{x}^{(k)})\, \nabla_{\underline{x}^{(k)}} P(\underline{x}^{(k)})^T$$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k\, \underline{S}^k$$

$$\boxed{\begin{array}{l} \underline{S}^k = -\underline{\underline{H}}^{-1}(\underline{x}^{(k)})\, \nabla_{\underline{x}^{(k)}} P(\underline{x}^{(k)})^T \\ \lambda_k = 1 \rightarrow \text{"Full step Newton's method"} \end{array}}$$

→ if $\underline{\underline{H}}^{-1}(\underline{x}^{(k)})$ is $\underline{\underline{I}}$ or diagonal matrix with same elements.

$\underline{\underline{H}}$ if not quadratic it changes, you must find Eigen values, if not +ve definite it diverge

solve it ☹

→ Example

$$\text{Min } p(\underline{x}) = 4x_1^2 + x_2^2 - 2x_1 x_1$$

use $x_0 = [1 \quad 1]^T$

ⓐ analytical method

$$\nabla p(\underline{x}) = 0^T$$

$$\frac{dP}{dx_1} = 0 \rightarrow 8x_1 - 2x_2 = 0 \qquad x^* = [0 \quad 0]^T$$

$$\frac{dP}{dx_2} = 0 \rightarrow 2x_2 - 2x_1 = 0$$

ممكن الاشتقاق يكون order أكبر بالتالي تحتاج

إلى numirical من أجل إيجاد الجذور.

ⓑ Newton's method

$$\nabla p(\underline{x})^T = \begin{bmatrix} 8x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{bmatrix} \qquad \underline{\underline{H}}(\underline{x}) = \begin{bmatrix} 8 & -2 \\ -2 & 2 \end{bmatrix}$$

cofactor matrix

Inverse ① $\underline{\underline{H}}^c(\underline{x}) = (-1)^{i+j} M_{ij}$ — same order of $\underline{\underline{H}}$

M_{ij} : determinant of the matrix ~~that~~ of $n-1$ order which excludes row i & column j.

$$\underline{\underline{H}}^c(\underline{x}) = \begin{bmatrix} 2 & +2 \\ 2 & 8 \end{bmatrix}$$

adjoint matrix

$$② \quad \underline{\underline{H}}^a = \underline{\underline{H}}^{c\,T} = \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix}$$

③ $\underline{\underline{H}}^{-1} = \frac{1}{|\underline{\underline{H}}|} \underline{\underline{H}}^{a} = \frac{1}{12} \underline{\underline{H}}^{a} = \boxed{\begin{bmatrix} 1/6 & 1/6 \\ 1/6 & 2/3 \end{bmatrix}}$

$\rightarrow \nabla p(\underline{x}^{(k)})^T = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$

$\rightarrow \underline{S}^{(0)} = -\underline{\underline{H}}^{-1}(\underline{x}^{*}) \nabla^T p(x)_{x^{(0)}} = -\begin{bmatrix} 1/6 & 1/6 \\ 1/6 & 2/3 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \end{bmatrix} = -\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\rightarrow \underline{x}^{(1)} = \underline{x}^{(0)} + 1\,\underline{S}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ because it is quadratic.

$\rightarrow$ if $p(\underline{x}) = 4x_1^3 + x_2^2 - 2x_1x_2$ it will differs !

$x_1 = 0, \frac{1}{6} \rightarrow x_2 = 0, \frac{1}{6}$ $\quad \underline{x}^{*} = [0 \;\; 0]^T$

$= [1/6 \;\; 1/6]^T$

check on $\underline{\underline{H}}$ if this function has optima.

$\underline{\underline{H}}(\underline{x}) = \begin{bmatrix} 24 & \\ & \end{bmatrix}$ $\quad \underline{\underline{H}}(x^{(0)}) = \begin{bmatrix} 0-\lambda & -2 \\ -2 & 2-\lambda \end{bmatrix}$ Find.

Note: if $p(\underline{x})$ is highly non linear, the Newton's method may diverge when the $\underline{\underline{H}}(\underline{x}^{(k)})$ is not positive definite (or semi definite).

check on $\underline{\underline{H}}$; if 0 or +ve no problem

problem if -ve $\rightarrow$ my initial guess is not feasible

To solve this problem $\rightarrow$ use another initial guess

In such case, enforce the $\underline{\underline{H}}$ to be +ve definite

(in any iteration, I want +ve $\underline{\underline{H}}$ if minimization).

$\rightarrow$ How to have +ve definite ?

اضافة على عناصر ال diagonal قيمة كبيرة +ve

$\tilde{\underline{\underline{H}}} = \underline{\underline{H}} + \beta\, \underline{\underline{I}}$

large scalar value to guarantee +ve definite matrix $\tilde{\underline{\underline{H}}}$

اذا ضفنا فوق قيمة كبيرة، مش كثير $\beta = 10$ بيكون حسب السؤال، اذا كبيرة جدا بتوصل نقطة ال optimum $\lambda_k \neq 1$ بعمل فحص شرطية

if loop statement in excel

Excersice

check on Two points, Find λ_1 & λ_2, if +ve يعني ما بنزيد إشي

** Quasi - Newtons method

inverse من المعروف

$$\underline{\underline{H}}_{\underline{x}^{(k)}}\left(\underline{x}^{(k+1)}-\underline{x}^{(k)}\right)=\nabla_{\underline{x}^{(k+1)}}p(\underline{x})^T-\nabla_{\underline{x}^{(k)}}p(\underline{x})^T \quad \frac{d^2f}{dx^2}=\frac{d}{dx}\left(\frac{df}{dx}\right)$$

" Forward finite difference method "

$$FD: f''(x_i)=\frac{\Delta}{x_{i+1}-x_i}\left(\frac{f_{i+1}-f_i}{x_{i+1}-x_i}\right)$$

~~→ assume initial guess direction since you don't have $\nabla_{\underline{x}^{(k+1)}} p(\underline{x})$~~

→ The Broyden, Fletcher, Goldforlg and Shanno (BFGS)

- method to approximate $\underline{\underline{H}}_{x^{(k+1)}}$:

$$\underline{\underline{H}}_{k+1}=\underline{\underline{H}}_k+\frac{\Delta g_k\,\Delta g_k^T}{\Delta g_k^T\,\Delta\underline{X}}-\frac{\underline{\underline{H}}_k\,\Delta\underline{X}\;\Delta\underline{X}^T\,\underline{\underline{H}}_k}{\Delta X^T\,\underline{\underline{H}}_k\,\Delta\underline{X}}$$

where $\Delta g_k=\nabla_{k+1} p(X)^T-\nabla_k p(\underline{X})^T$

$$\Delta X = X_{k+1}-X_k$$

Scanned by CamScanner

* Approach of numerical methods :

① select an initial guess point(s)

② Find the direction which improves the value of $p(x)$

③ move along that direction

④ repeat ② and ③ until converged

→ The different numerical methods come from the way in which it uses the search direction.

25/2

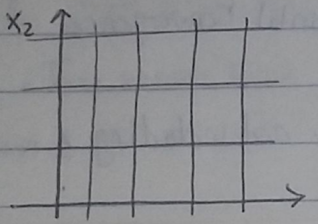
* Numerical methods for unconstrained multi-variable optimization function

① Random search method

min $p(\underline{x})$

X_1	X_2	X_3	$\cdots X_n$	$P(\underline{X})$

mesh (grid)



100 point with 10 variables

↳ calculation of $p(x)$

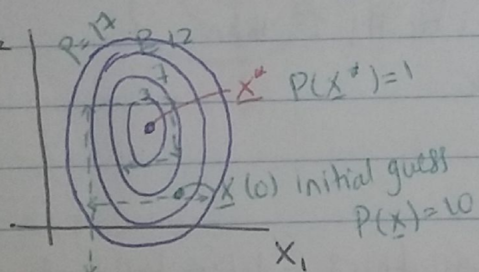
→ Huge calculation of $p(x)$ when n is large (pos but it can be).

② Univariate search method

→ search for the optimum in each direction independently

نثبت كل شيء و
بنغير واحد فقط

Min $p(\underline{x})$



سنلاحظ $p(x)$ تزيد و توقف
و نروح للمتغير الآخر و هكذا
"orthoganity)"

→ effective for $p(\underline{x})$ of the form $p(x) = \sum_{i=1}^{n} c_i x_i^2 \rightarrow \underline{x}^* = 0^T$

if $x_1, x_2 \rightarrow$ circles

$x_1, x_2, x_3 \rightarrow$ balls

→ less effective for $p(\underline{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_i x_j$ ellipse

~~constant + x_1 + $x_1 x_2$~~ $\quad Ax_1^2 + Bx_2^2 + C(x_1 x_2)$

Most of software

② Simplex method (in matlab)

- Simplex : geometric figure for by n+1 points in the n-dimensional space   no orthogonality

if $x_1, x_2 \rightarrow$ 3 points $\rightarrow$ triangle

P_2

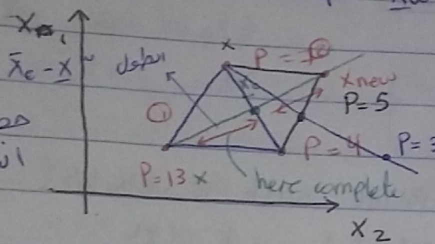
P_3 P_1 (initial guess)   I know were is the worst point.

Procedure

1. start with a simplex of n+1 vertices, where $\dim(\underline{x}) = n$
2. evaluate $P(\underline{x})$ at each vertex and find the worst point ($\underline{x}_w$) (which has the highest $p(\underline{x})$)
3. reflect the x_w through the centroid of the simplex formed by all the points except $\underline{x}_w$.   middle between the other two points

بالتجريب

حجم المثلث يصغر كلما اقتربنا



Reflection — complete
expansion   أنت بالتجريب
contraction

And so on until convergence of x_{new} with x_{old}

4. Form the new simplex by rejecting the $\underline{x}_w$ and including $\underline{x}_{new}$
5. repeat 2-4 untill converged.

$$\boxed{\underline{x}_c = \frac{\sum_{i=1}^{n+1} \underline{x}_i - \underline{x}_w}{n}}$$

$$\boxed{\underline{x}_{new} = \underline{x}_c + \alpha(\underline{x}_c - \underline{x}_w)}$$

if α = 1 complete reflection
depends on $p(x)$ $\alpha > 1$ reflection and expansion
< 1 reflection and contraction

only finding 1 f(p(x)) each iteration

- simplex method has slow rate of convergence
  - You make investigation on α, firstly 1, then iterations $\rightarrow$ better α value

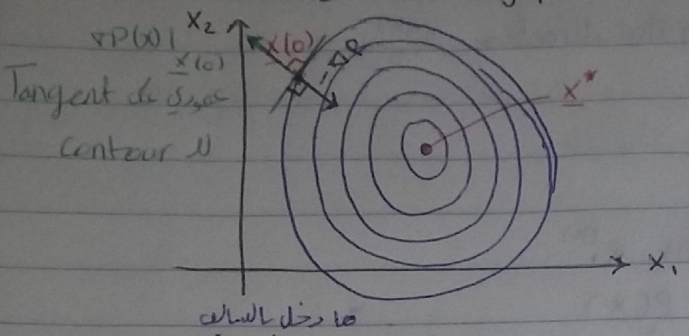
In Matlab fminsearch

Until know they don't depends on derivatives, let's take derivatives

** Gradient based method (it is raw vector)

$\nabla p(\underline{x})$

→ The direction of $\nabla p(\underline{x})$ is orthogonal (perpendicular) to the $p(\underline{x})$ contours at any point.



→ a descent direction $\underline{s}$ which direct to minima must satisfy

dot product يرجع الاتجاه

$\nabla p(\underline{x}) \cdot \underline{s} \leq 0$ if $\underline{s} = -\nabla p(\underline{x})$

$\nabla p(x)\ s \cos\theta$

هنا $180°$ المحصورة بين s و ∇p $90 < \theta \leq 180$

اذا اقل من $90°$ سيء

• This means that if $\theta = 90°$: no improvement in optima.

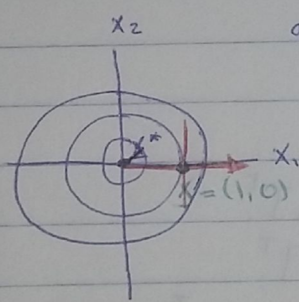
if $0 \leq \theta \leq 90°$: no improvement

if $90 < \theta \leq 180°$: improvement of optima (minima).

<u>Example</u> Min $p(\underline{x}) = x_1^2 + x_2^2$

$x^* = [0 \quad 0]^T$ $\nabla P(\underline{x}) = [2x_1 \quad 2x_2]^T$

at $\underline{x}^{(0)} = [1, 0]^T$: $\nabla P = [2 \quad 0]$ خارج منه



even at $[1 \quad 1]^T$, gradient is $\perp$ tangent

set $k = 0$

$$\underline{x}^{(k+1)} = \underline{x}^{(k)} + \underbrace{\lambda_k}_{\text{مقدار}} \underbrace{\underline{s}^{(k)}}_{\text{اتجاه}}$$

(need optimum value)

line search

ويعين متغير اتجاه اذا تطلب

Wednesday
27/2/2019

for minimization (maximization → ascent)

** steepest descent method :

→ initial guess point, $\underline{x}^{(0)}$ of $\underline{x}^*$ for minimization.

→ The search direction : $\underline{s}^{(k)} = (-) \nabla_{\underline{x}^{(k)}} P(\underline{x})$

→ Rule : $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \, \underline{s}^{(k)}$ $k = 0, 1, 2, \ldots$

λ_k : optimum step size along the direction $\underline{s}^{(k)}$

Algorithm

1. set $k = 0$
2. select initial guess point $\underline{x}^{(k)}$
3. determine $\underline{s}^{(k)} = -\nabla_{\underline{x}_k} P(\underline{x})$
4. perform line search to obtain λ_k

على اتجاه واحد — $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$ only λ_k is unknown

unidirection — $p(\underline{x}) = \ldots$

↓ single variable optimization — $p(\underline{x}^{(k+1)}) = G(\lambda_k)$

$$\frac{dP(\underline{x}^{(k+1)})}{d\lambda_k} = 0 \rightarrow \text{Find } \lambda_k^*$$

5. evaluate $\underline{x}^{(k+1)} = \underline{x}^{(k)} + \lambda_k \underline{s}^{(k)}$
6. If $\sum_{i=1}^{n} |x_i^{(k+1)} - x_i^{(k)}| / |x_i^{(k)}| < \epsilon$ stop — Convergence criteria on P or X
7. $k = k+1$
8. go to step 3

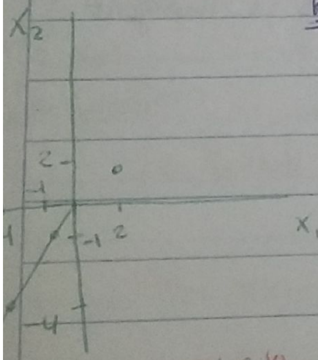
Example

$\min p(\underline{x}) = x_1^2 + x_2^2$

$k=0$ — $\underline{x}^{(0)} = [2 \quad 2]^T$

$\nabla p(\underline{x}) = [2x_1 \quad 2x_2]^T$

$\nabla_{\underline{x}(0)} P(\underline{x}) = [4 \quad 4]^T$



$\underline{s}^{(0)} = -\nabla_{\underline{x}(0)} P(\underline{x}) = [-4 \quad -4]^T$ You can take $[-1 \quad -1]^T$ because it is a vector

if you multiply or divide a vector by scalar

الأفضل أن لا تتغير الإشارة، عشان نخلي λ موجبة، ما بطلع عكس الاتجاه

$$\underline{x}^{(1)} = \underline{x}^{(0)} + \lambda_0 \underline{s}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \lambda_0 \begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 2-4\lambda_0 \\ 2-4\lambda_0 \end{bmatrix}$$

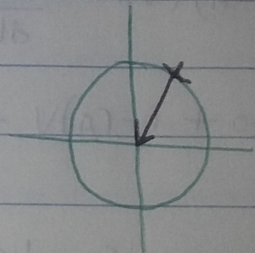
Line search : $p(\underline{x}^{(1)}) = p(x_1^{(1)}, x_2^{(1)}) =$

$p(2-4\lambda_0, 2-4\lambda_0)$

$$P(\underline{x}_1^{(0)}) = (2-4\lambda_0)^2 + (2-4\lambda_0)^2 = 2(2-4\lambda_0)^2$$

$$\frac{dP}{d\lambda_0} = 0 \rightarrow 4 * -4(2-4\lambda_0) = 0 \qquad 2-4\lambda_0 = 0 \quad \boxed{\lambda_0 = \frac{1}{2}}$$

substitute in $\underline{x}^{(1)} = \begin{bmatrix} 2-4(0.5) \\ 2-4(0.5) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



المفروض انه ما نوقف هون اذا بنقدر نحل analytical
ما نستخدم هاي الطريقة.

Quiz

$A \rightarrow B$ $\quad\quad I - O + G - C = A$

$B \rightarrow C$

(A) $-(-r_A)(V) = \frac{dN_A}{dt}$ $\quad\quad -(r_A) = \frac{dC_A}{dt}$

$0 - 0 + (-r_A)V - (r_c)V = \frac{dN_B}{dt}$

$-r_A - r_c = \frac{dC_B}{dt}$ $\quad\quad r_A = k_1 C_A$

non homogeneous D.E

Linear Programming LP

Wednesday 13/3/2019

* iterations are required when linearization. (e.g: secant method)

*** why?**

- Some optimization problem are linear or can be approximated by linear forms.
- Frequently used in optimization.
- well-developed thiories for LP
- easy to obtain sensitivity information (analysis)
- help to understand the non-linear programming problems.
- excellent soft-wares are available as LINDO

*** General formulation (not standard) of LP problem**

constrained LP is of interest

equality constrains are linear

$$\begin{matrix}\text{Min}\\\text{Max}\end{matrix} \quad p(\underline{x}) = \underline{c}^T \underline{x} \qquad \underline{c}^T = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} : \text{constant coellicients}$$

$$\text{s.t} \quad \underline{\underline{A}}^{eq} \underline{x} = b^{eq.}$$

$$\underline{\underline{A}}^{ineq} \underline{x} \geq b^{ineq}$$

$$x_{min} \leq \underline{x} \leq x_{max}$$

→ Graphical (solution) can't be ~~x~~ $n > 3$ no more than 3D

representation

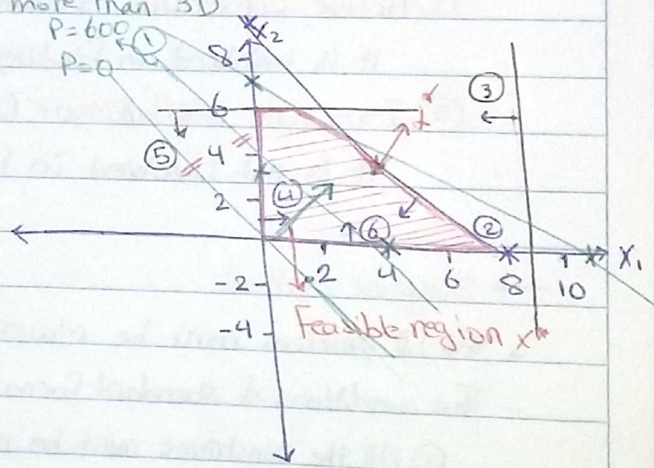
Example $\max p(\underline{x}) = 150x_1 + 175x_2$

s.t $7x_1 + 11x_2 \leq 77$ ①

$10x_1 + 8x_2 \leq 80$ ②

$x_1 \leq 9$ ③ $x_1 \geq 0$ ④

$x_2 \leq 6$ ⑤ $x_2 \geq 0$ ⑥



① $7x_1 + 11x_2 = 77$ $\quad x_2 = 7 - \frac{7}{11}x_1$

② $10x_1 + 8x_2 = 80$ $\quad x_2 = 10 - \frac{\cancel{80}10}{\cancel{10}8}x_1$

* levels of $P(x)$ since you want max., increase $p(x)$

slope does not change

ⓐ $P = 0$ $\quad 0 = 150x_1 + 175x_2 \quad x_2 = \frac{-150}{175}x_1 = -0.857x_1 + 0$

ⓐ $P = 6$ $\quad 6 = 150x_1 + 175x_2 \quad x_2 = -0.857x_1 + \frac{6}{175}$ → hard to draw

ⓐ $P = 600$ $\quad x_2 = -0.857x_1 + \frac{600}{175} = -0.857x_1 + 3.42$

→ intersection between ①&② , the line slope is not 1.25 nor -0.86 but between

$7x_1 + 11x_2 = 77$ $\quad$ → $\underline{x}^* = [4.9 \quad 3.9]^T$

$10x_1 + 8x_2 = 80$ $\quad P^*(\underline{x}^*) = \cancel{70}1417.5$

if $p(\underline{x}) = 150x_1$ → optimum at $x_1 = 8$ $x_2 = 0$ vertical lines

$p(\underline{x}) = 175x_2$ → optimum at $x_1 = 0$ $x_2 = 6$ horizontal lines.

$p(\underline{x}) = 150x_1 + 300x_2$ at ~~corner~~ corner

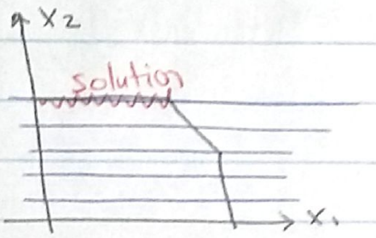
always, LP the optima is على الدوام بتطلع

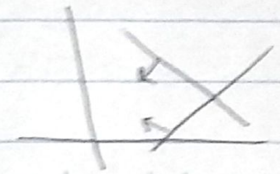
النقطة المثالية اذا مثلاً الحدود افقية، ال counters أفقية.

Imp. Remark The optimum point is located at one vertix (corner point) of the feasible region for well-posed LP problem. → all bounded

→ Simplex method is based on this fact.

Exception multiple solution (Degeneracy)





unbounded region

No closed feasible region

"improper formulation"

→ Constrains are divided

① Active constrains (bound)

it is involved in finding $\underline{x}^*$ (like

② Inactive constrains (or bound)

it is not involved in finding $\underline{x}^*$

→ Simplex method

- The LP problem must be rewritten in a standard form for simplex method
- The conditions of standard form

① All the constrains must be equations with non-negative (+ve) RHS (right hand side) coefficients

② All variables must be non-negative (≥ 0)

③ The $p(x)$ can be max. or min.

* Some helpful rules :-

① For inequality constrains with $\leq$:

Convert this inequality by adding slack (ركود) variable ($S \geq 0$)

Example $x_1 + x_2 \leq 6$

$x_1 + x_2 + S_1 = 6$

if the constrains represents the limit of usage of resource, S will be the unused amount of such resource.

1 2 7

heat
mass

② For inequality constrains with $\geq$:

convert this inequality by substracting surplus variable ($S \geq 0$) to have all variable +ve.

Example $x_1 + 3x_2 + 7x_3 \geq 0$

$x_1 + 3x_2 + 7x_3 - S = 0$

dis .... RHS

$+x_1 + 3x_2 + 7x_3 \geq -4$ $\times -1$

$-x_1 - 3x_2 - 7x_3 \leq 4$ use slap

Firstly → ③ multiply the constrain by -1 if its right hand side is negative — Monday 18/

take into consideration the $\leq$ become $\geq$

$\geq$ become $\leq$

④ express the unrestricted variable x_i as a dilference of two positive variables

$x_i = x_i^* - x_i^{**}$ such that $x_i^* \geq 0$ $x_i^{**} \geq 0$ or any named variable

x_i' x_i'' not derivatives

Example write the following LP problem in a standard form for simplex method

min $p(\underline{x}) = 2x_1 + 3x_2$

s.t : $x_1 + x_2 = 10$

$-2x_1 + 3x_2 \leq -5 \rightarrow 2x_1 - 3x_2 \geq 5$ $2x_1 - 3x_2 - S_1 = 5$

$7x_1 - 4x_2 \leq 6 \rightarrow 7x_1 - 4x_2 + S_2 = 6$

$x_2 \geq 0$

x_1 unrestricted $x_1 = x_1' - x_1''$ $x_1' \geq 0$ $x_1'' \geq 0$

- Our problem becomes

min $p(x) = 2x_1' - 2x_1'' + 3x_2$

s.t: $2x_1' - 2x_1'' - 3x_2 - S_1 = 5$

$7x_1' - 7x_1'' - 4x_2 + S_2 = 6$

$x_1' - x_1'' + x_2 = 10$

$x_1', x_1'', x_2, S_1, S_2 \geq 0$

Sheet example

max profit $= p(\underline{x}) = x_1 + 3x_2$ x_1: lbm of A x_2: lbm of B

↳ min $-p(\underline{x}) = -x_1 - 3x_2$

s.t: sugar ≤ 6 lbm $0.2x_1 + 0.1x_2 \leq 6$

chocolate ≤ 5.5 lbm $0.1x_1 + 0.2x_2 \leq 5.5$

milk ≤ 10 lbm $0.1x_1 + 0.4x_2 \leq 10$

$x_1 \geq 0$

$x_2 \geq 0$

always initial guess is 0,0 in simplex method

Standard form $\quad \text{Min} \quad -x_1 - 3x_2 + 0s_1 + 0s_2 + 0s_3$

s.t $\quad$ sugar $\quad 0.2x_1 + 0.1x_2 + s_1 = 6 \qquad$ if $s_1 = 0 \rightarrow$ كل الحد راح

choc. $\quad 0.1x_1 + 0.2x_2 + s_2 = 5.5$

milk $\quad 0.1x_1 + 0.4x_2 + s_3 = 10$

$x_1, x_2, s_1, s_2, s_3 \geq 0$

equations = m = 3

$$\min \underset{m\times n}{\underline{c}^T} \underset{n\times 1}{\underline{x}}$$

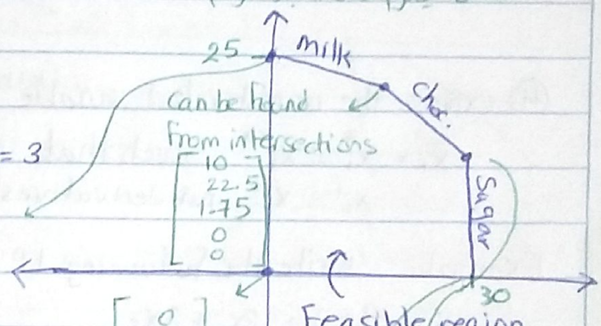
$$\text{s.t } \underline{A}\,\underline{x} = \underset{m\times 1}{\underline{b}} \quad \text{where } \underline{b} \geq 0 \qquad \underline{x} \geq 0$$

s لا يلزم من أجل الرسم

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} n=5 \qquad \underline{b} = \begin{bmatrix} 6 \\ 5.5 \\ 10 \end{bmatrix} m=3$$

equality constrains

$$\begin{bmatrix} 0 \\ 25 \\ 3.5 \\ 0.5 \\ 0 \end{bmatrix}$$



$$\begin{bmatrix} 10 \\ 22.5 \\ 1.75 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 6 \\ 5.5 \\ 10 \end{bmatrix} \quad \begin{bmatrix} 30 \\ 0 \\ 0 \\ 2.5 \\ 7 \end{bmatrix} \quad \begin{bmatrix} 21.67 \\ 16.67 \\ 0 \\ 0 \\ 1.167 \end{bmatrix}$$

$$\underline{A} = \begin{bmatrix} 0.2 & 0.1 & 1 & 0 & 0 \\ 0.1 & 0.2 & 0 & 1 & 0 \\ 0.1 & 0.4 & 0 & 0 & 1 \end{bmatrix}$$

- there is two zeros in each point $= n - m = 5 - 3 = 2$ ※
- numbers become zero and numbers become non zero but still 2 zeros

Remark

1. always there are n-m zeros "non basic variables"
2. and they are m non zeros "basic variables"

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

x_1, x_2 — non basic

x_3, x_4, x_5, x_6 — basic (4)

leaving variable

entering variable

* Steps of simplex method

step 0 • determine a starting basic feasible solution by setting $(n-m)$ non basic variables at zero level. (select the slack variables as initial basics) (start from origin point as an initial guess)

step 1 • select an entering variable from among the current non-basic variables such that this entering variable will improve the value of $p(x)$

-3 أكبر من -20

max	$x_1 + 3x_2$	choose x_2	$5x_1 - 7x_2 + x_3$	choose x_1
min	$-x_1 - 3x_2$	choose x_2	if min $\to x_2$	

step 2 • select leaving variable from among m basic variable such that it will lead to maximum feasible value of the entering variable selected in step 1.

step 3 • Find the new basic solution

step 4 • go to step 1 until there is no further improvement

مثلاً objective جميع coefficient سوالب وهو max.

Simplex method - Tableau

→ objective function : p-equation $P(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n$

$P - c_1x_1 - c_2x_2 - \dots - c_nx_n = 0$

$n-m$: non-basic (zeros)

m : basic

including slack

Basic	P	x_1 x_2 x_3 --- x_n	~~right~~ RHS
P	1	$-c_1$ $-c_2$ $-c_3$ .... $-c_n$	0
x_1	0		
x_2	0	$\underline{\underline{A}}$	$\underline{b}$
⋮	⋮		
x_n	0		

initial slack

max $3x_2 + x_1$

$P - 3x_2 - x_1 = 0$

if min → +ve in optimize

أكبر شي بالموجب

أكبر شي بالسالب → max

* steps of simplex method

step 0 • determine a starting basic feasible solution by setting $(n-m)$ non basic variables at zero level. (select the slack variables as initial basics) (start from origin point as an initial guess)

step 1 • select an entering variable from among the current non-basic variables such that this entering variable will improve the value of

-3 أكبر من -20

$p(x)$ — max $x_1 + 3x_2$ choose x_2 — $5x_1 - 7x_2 + x_3$ choose x_1

min $-x_1 - 3x_2$ choose x_2 — if min → x_2

step 2 • select leaving variable from among m basic variable such that it will lead to maximum feasible value of the entering variable selected in step 1.

step 3 • find the new basic solution

step 4 • go to step 1 until there is no further improvement

يعني objective دجميع coefficient سوالب , dمو max.

Simplex method - Tableau

→ objective function : p-equation $p(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n$

$P - c_1x_1 - c_2x_2 - \dots - c_nx_n = 0$

$n-m$: non-basic (zeros)

m : basic

including slack

Basic	P	x_1	x_2	x_3 ---	x_n	~~right~~ RHS
P	1	$-c_1$	$-c_2$	$-c_3$	$-c_n$	0
x_1	0					
x_2	0		$\underline{\underline{A}}$			$\underline{b}$
x_n	0					

initial slack

ممنوع تكون negative

max $3x_2 + x_1$

$P \boxed{-3x_2} - x_1 = 0$

if min → +ve in optimize

أكبر شي بالموجب

أكبر شي بالسالب → max

Monday 25/3/2018

step 0 initial feasible solution

→ select the slack variables as basic

Basic	P	x_1	x_2	* s_1	s_2	s_3	RHS	RHS/a_i
P	1	-1	-3	0	0	0	0	
s_1	0	0.2	0.1	1	0	0	6	60
s_2	0	0.1	0.2	0	1	0	5.5	27.5
s_3	0	0.1	0.4	0	0	1	10	25

become x_2 new solution is required

"pivot equation" "pivot element" leaving mid exam review

any non linear constrain → NLP
actual error → between exact and iteration
approximate error → between iterations
3 iterations → without initial guesses
8 تبلغ ي decimals انا العبدكبير
F(x) → quadatic quadratic approximation becomes not approximation
↳ answer from the first iteration of Newton's method

Step 1 select entering variable in a way that improve the value of o.f

X_2 is entering variable

Step 2 select leaving variable from the basic set such that the entering variable will be within the feasible region

befor step 2, $S_1 = 6$ (since $x_1 = 0$ $x_2 = 0$)

$S_2 = 5.5$

I want to make S_1 or S_2 or S_3 equal zero but X_2 still in the feasible region

(X_1 still zero) if $S_1 = 0$ $\quad 0.1 X_2 = 6$ $\quad X_2 = 60$ is it in the feasible region? no

if $S_2 = 0$ $\quad 0.2 X_2 = 5.5$ $\quad X_2 = 27.5 \rightarrow$ non feasible

if $S_3 = 0$ $\quad 0.4 X_2 = 10$ $\quad X_2 = 25 \rightarrow$ feasible (in the band (vertix))

* So ... ratio test is needed instead of trying

ratio $= \frac{RHS}{a_i}$ the coefficient of row i in the entering variable column.

Then select the row with the smallest ratio to select the leaving variable

لما أنه مسموح تاخد الصفر (بيصير ... ونقرب ...) ✓

Step 3 Determine the new basic solution: Use Gaussian-Jordan method

→ Divide the pivot equation by the value of pivot element

· ~~The~~ last row becomes (R_4)

X_2 0   0.25   1   0   0   2.5   25 $RHS/a_i \rightarrow$ no meaning now

in order to main the coefficients of the entering variable = zero

→ Perform mathmatical operations such all the coefficients of the entering variable column are zero except the pivot element (coefficient = 1)

* replace row i by: $R_i = R_i - a_i * \text{pivot equation}$

→ $R_1^* = R_1 - (-3) R_4$

→ $R_2^* = R_2 - (0.1) R_4$

→ $R_3^* = R_3 - (0.2) R_4$

all coefficients will be changed or the Rows except R_4

	P	X_1	X_2	S_1	S_2	S_3	RHS	X_1 still nonbasic
P	1	-0.25	0	0	0	7.5	75	$X_1 = 0$
S_1	0	0.175	0	1	0	-0.25	3.5 ←	$S_1 = 3.5$
S_2	0	0.05	0	0	1	-0.5	0.5	
X_2	0	0.25	1	0	0	2.5	25	

* RHS are representing non basic values ✱

actually it is non zero * zero or zero * non zero

variable   coefficient

still there is optimization -ve coefficient since max. p(x) { min → +ve take أكبر ; max → -ve أصغر }

X_1 is the entering variable

x_1 is the entering variable , make ratio test to choose leaving variable

RHS / a_i

20

10 → min. → take S_2 pivot equation in this raw → 0 | 1 | 0 | 0 | 20 | -10 | 10

100

Remark always RHS must be +ve, since it is the basic variables solution
if it becomes -ve , multiply by -1

$R_1^{**} = R_1^* - (-0.25) R_3$

$R_2^{**} = R_2^* - (0.175) R_3$

$R_4^* = R_4 - (0.25) R_3$

P	1	0	0	0	5	~~7~~5	77.5	
S_1	0	0	0	1	-3.5	1.5	1.75	~~S_1 = 1.75~~
X_1	0	1	0	0	20	-10	10	
X_2	0	0	1	0	0	5	22.5	

no further optimization   The End ∇

no further improvement of o.f

$P^* = 77.5$

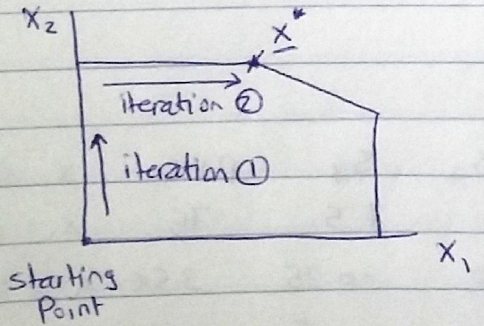
$X_1^* = 10$

$X_2^* = 22.5$

$S_1^* = 1.75$ →   كمية ما ذخرته من السكر   amount المستهلكة = 6-1.75

$S_2^* = 0$ → (not in basic)

$S_3^* = 0$



قام بسلك الطريق الأقصر

You can evaluate p at each vertix → find $\underline{X}$

Wednesday
27/3/2019

(initial step 0 is important)

* Initial feasible solution

→ Select the slack variables as initial basics of non-zero values

→ # of basics = m = # of equations in the standard form

n = # of variables in the standard form

$n - m$ = # of basics with zero values.

→ For any equation without slack variable, add what is called artificial variable ($R_i \geq 0$) and then select this R_i to be in the initial basic set of variables.

(artificial variable method)

Example
it can be drawn since 2 variables

min $p = 4x_1 + x_2$

s.t : $3x_1 + x_2 = 3$

$4x_1 + 3x_2 \geq 6$

$x_1 + 2x_2 \leq 4$

$x_1, x_2 \geq 0$

if not given in question, you can not assume it

اذا كان سالب اضرب ب -1 مشان تحصل slack ؟

standard form : min $p = 4x_1 + x_2$

s.t ~~$4x_1 + 3x_2 - x_3 + R$~~

$3x_1 + x_2 + R_1 = 3$

$4x_1 + 3x_2 - x_3 + R_2 = 6$

$x_1 + 2x_2 + x_4 = 4$

$x_1, x_2, x_3, x_4, R_1, R_2 \geq 0$

$m = 3$ Basic (initial)

$n = 6$

$m = 3$ basic

$n - m = 3$ non basic

** There are two methods deal with the artificial formulation in order to get the final optimum solution

(A) Two-phase method (B) Big-M technique

(A) phase I
optimization of artificial variables

min $r = R_1 + R_2$ ← no -ve, make -ve terms from the below equations

s.t $3x_1 + x_2 + R_1 = 3 \rightarrow R_1 = 3 - 3x_1 - x_2$

$4x_1 + 3x_2 - x_3 + R_2 = 6 \rightarrow R_2 = 6 - 4x_1 - 3x_2 + x_3$ } substitute in r

$x_1 + 2x_2 + x_4 = 4$

$x_1, x_2, x_3, x_4, R_1, R_2 \geq 0$

$r = 3 - 3x_1 - x_2 + 6 - 4x_1 - 3x_2 + x_3 = 9 - 7x_1 - 4x_2 + x_3$

r eq. : $r + 7x_1 + 4x_2 - x_3 = 9$

Then Table

Basic	r	x_1 ↓	x_2	x_3	x_4	R_1	R_2	RHS
r	1	7	4	-1	0	0	0	9
→ R_1	0	3	1	0	0	1	0	3
R_2	0	4	3	-1	0	0	1	6
x_4	0	1	2	0	1	0	0	4

← Stop when all coefficients are -ve

choose according to test (leaving variable)

Entering variable : x_1

leaving variable : R_1

Basic	r	x_1	x_2	x_3	R_1 ~~x_4~~	R_2 ~~R_1~~	x_4 ~~R_2~~	RHS
r	1	0	0	0	-1	-1	0	0
x_1	0	1	0	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{-1}{5}$	0	$3/5$
x_2	0	0	1	$\frac{-3}{5}$	$\frac{-4}{5}$	$3/5$	0	$6/5$
x_4	0	0	0	1	1	-1	1	1

↑ at optimum $r^* = 0$ (غير موجود في العمود الأول)

$R_1^* = 0$

$R_2^* = 0$ As required ! (good as initial guess for Phase II)

Phase II

- Use the optimum (basic) solution of phase I as a starting initial solution (choice) for the original problem.

take equations from Table ← تقلل من المتغيرات

$$\min\ p(\underline{x}) = 4x_1 + x_2$$

s.t

$$x_1 + 1/5\, x_3 = 3/5$$

$$x_2 - 3/5\, x_3 = 6/5$$

$$x_3 + x_4 = 1$$

$$x_1,\ x_2,\ x_3,\ x_4 \geq 0 \qquad n = 4 \quad m = 3$$

non basic → 1 variable.

- But $p(x)$ coefficients are all +ve, min I want 1 -ve coefficient

$x_1 = 3/5 - 1/5\, x_3 \qquad x_2 = 6/5 + 3/5\, x_3$ substitute in P

$$P(\underline{x}) = 4\left(\frac{3}{5} - \frac{1}{5}x_3\right) + \left(\frac{6}{5} + \frac{3}{5}x_3\right) = \frac{-1}{5}x_3 + \frac{18}{5}$$

P equation : $P + \frac{1}{5}x_3 = \frac{18}{5}$ $\quad x_3$ is the only +ve variable (non basic variable) standard from ~~require~~ includes artificial variables

Basic	P	X_1	X_2	X_3	X_4	RHS
P	1	0	0	1/5	0	18/5
X_1	0	1	0	1/5	0	3/5
X_2	0	0	1	-3/5	0	6/5
→ X_4	0	0	0	1	1	1

↖ Initial

?

↑

Basic	P	X_1	X_2	X_4	RHS
P	1	0	0	-1/5	17/5
X_1	0	1	0	-1/5	2/5
X_2	0	0	1	3/5	9/5
X_3	0	0	0	1	1

$P^* = 17/5$

$X_1^* = 2/5$ } You can check by substituting in $P(\underline{X})$

$X_2^* = 9/5$

$X_3^* = 1$

$X_4^* = 0$ ⇝ all 4 are gone

if all problem is slack, all -ve coefficients ↗ This is the easiest way

(B) M : very large positive number (we are not interested in its value) $M > 0$

→ Incorporate the artificial variables in the original objective function by penalizing them with very big number (M).

- How?

- For minimization problem

$$\text{Min } C^T \underline{X} + M \sum_{i=1}^{\#\text{ of artificial}} R_i$$

trend of decreasing

- For maximization problem

$$\text{Max } \underline{C}^T \underline{X} - M \sum R_i$$

trend of increasing

if +ve it will stop from the initial

- artificial variables in terms of other variables

Previous example Min $P = 4X_1 + X_2$

$P = 4X_1 + X_2 + MR_1 + MR_2$

s.t $3X_1 + X_2 + R_1 = 3$

$4X_1 + 3X_2 - X_3 + R_2 = 6$

$X_1 + 2X_2 + X_4 = 4$

$X_1, X_2, X_3, X_4, R_1, R_2 \geq 0$

$n = 6 \quad m = 3$

$R_1 = 3 - 3X_1 - X_2$

$R_2 = 6 - 4X_2 - 3X_2 + X_3$

P eq. : $P - 4X_1 - X_2 + -M(3 - 3X_1 - X_2) - M(6 - 4X_1 - 3X_2 + X_3) = 0$

$P + (-4 + 7M)X_1 + (-1 + 4M)X_2 + ^{-}M X_3 = 9M$ look at M

entering variable : X_1 ~~Basics R_1 R_2 X_2~~

starting basics : R_1 R_2 X_4

Return to book for the solution.

solve all previously solved examples.

Lindo program

make test & on feasibility

1/4/2019

- Standard form is not required , slack + $P(\underline{x}^*) + \underline{x}^*$ are of interest

→ sentivity analysis

- no model is perfect

→ How will the optimum solution change when the coefficients [$\underline{c}$, $\underline{b}$, $\underline{\underline{A}}^{eq}$, $\underline{\underline{A}}^{ineq}$] are changed.

original problem

** Primal - Dual problem

→ The dual problem is an axuiliary problem defined directly from the standard form of the primal problem (after taking artificial variables, for M-technique) into consideration.

→ primal : n variables (in the standard form) with slack, surplus, artificial
m equations → $x_1, x_2, \dots, x_n$

→ dual : m variables → $y_1, y_2, \dots, y_m$
n equations   Solving dual is easier than dual (if big m)

→ The dual of the dual is primal

Basics	P	x_1	$x_2 \cdots$	$x_j \cdots$	x_n	RHS	
P	1	$-c_1$	$-c_2 \cdots$	$-c_j \cdots$	$-c_n$	$-$(some value)	Primal
⋮	0	a_{11}	$a_{12} \cdots$	$a_{1j} \cdots$	a_{1n}	b_1	y_1
⋮	0	a_{21}	$a_{22} \cdots$	$a_{2j} \cdots$	a_{2n}	b_2	y_2
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	0	a_{m1}	$a_{m2} \cdots$	$a_{mj} \cdots$	a_{mn}	b_m	y_m

ما منقدر نحط المتغيرات لأنها مش معروف هون شو هي

* Rules to get the dual problem

p(x) 1 - If the primal talks about $\overset{\min}{\max}$ → the dual will be $\overset{\max}{\min}$ & vice versa

Primal: <u>max</u> $\underline{C}^T \underline{x}$

Dual : <u>min</u> $\underline{b}^T \underline{y} = b_1y_1 + b_2y_2 + \cdots + b_my_m$

2 - ~~If the~~ For primal problem of $\overset{\min}{\max}$ type, the n constrains of the dual problem will be of the type $\overset{\leq}{\geq}$ & vice versa   constrains

Primal : max $\underline{C}^T \underline{x}$ surplus & artificial variables while $\leq$ only slack

$a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m \geq c_1$ associated with x_1

$a_{12}y_2 + a_{22}y_2 + \cdots + a_{m2}y_m \geq c_2$ associated with x_2

⋮

$a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{nm}y_m \geq c_n$ associated with x_n

3- The variables of the dual problem are unrestricted.

ONLY RHS = 0 ~~& only in p(x)~~ $y \geq 0$
in constrains
no other variables in this constrain
(other coefficients are zeros)

3/4/2019

Example write the dual form of

min $Z = 5x_1 - 2x_2$

$-x_1 + x_2 \geq -3$

$2x_1 + 3x_2 \leq 5$

$x_1, x_2 \geq 0$

→ standard form   min $5x_1 - 2x_2 = Z$ $+0x_3 + 0x_4$ $n = 4$

Be carefull الترتيب

$x_1 - x_2 + x_3 = 3$ $m = 2$

$2x_1 + 3x_2 + x_4 = 5$ ↓

$x_1, x_2, x_3, x_4 \geq 0$ y_1, y_2 when dual

→ Dual   max $W = 3y_1 + 5y_2$ s.t

if you want to solve it, it is not in a standard form.

s.t	$y_1 + 2y_2 \leq 5$	associated with x_1
	$-y_1 + 3y_2 \leq -2$	" " x_2
	$y_1 \leq 0$	" " x_3
	$y_2 \leq 0$	" " x_4

Remark optimal solution of 1 problem is immediately obtained (without any computation of simplex) using the optimal simplex tableau of the other problem. Verify it !

Example max $Z = 5x_1 + 12x_2 + 4x_3$

s.t $x_1 + 2x_2 + x_3 \leq 10$

$2x_1 - x_2 + 3x_3 = 8$

$x_1, x_2, x_3 \geq 0$

$\xrightarrow{\text{standard form}}$ max $Z = 5x_1 + 12x_2 + 4x_3 + MR$

s.t y_1: $x_1 + 2x_2 + x_3 + x_4 = 10$

y_2: $2x_1 - x_2 + 3x_3 + R = 8$

$R, x_1, x_2, x_3, x_4 \geq 0$

$n = 5$ $m = 2$ – y_1, y_2

min $W = 10y_1 + 8y_2$

s.t x_1	$y_1 + 2y_2$	≥ 5
x_2	$2y_1 - y_2$	≥ 12
x_3	$y_1 + 3y_2$	≥ 4
x_4	y_1	≥ 0
R	y_2	$\geq \cancel{0}\ -M$

→ if you want to convert it to standard, don't convert this

&

y_2 → تبقى موجبة

Return to sheet for the primal solution

→ using the solution of primal, use the rules given below the table 4-3 (Page 122) The sheet.

$w^* = Z^* = 54\frac{4}{5}$

$\cancel{y_1}$, $y_1^* - 0 = \frac{29}{5}$ $\boxed{y_1^* = \frac{29}{5}}$ You can take it from the other constrains

$y_2^* - (-M) = \frac{-2}{5} + M$ $\boxed{y_2^* = \frac{-2}{5}}$

* Dual of the dual is primal

Try without cancelling R

Regenerate it

compare the Two phase and big M

Add z coefficients column.

Table 4-3

Iteration	Basic	x_1	x_2	x_3	x_4	R	Solution	
0 (starting)	z	$-5-2M$	$-12+M$	$-4-3M$	0	0	$-8M$	
x_3 enters	x_4	1	2	1	1	0	10	10
R leaves	R	2	-1	3	0	1	8	$\frac{8}{3}$
1	z	$-7/3$	$-40/3$	0	0	$\frac{4}{3}+M$	32/3	
x_2 enters	x_4	1/3	7/3	0	1	$-1/3$	22/3	* choose only the minimum value in the ratio test
x_4 leaves	x_3	2/3	$-1/3$	1	0	1/3	8/3	
2	z	$-3/7$	0	0	40/7	$-\frac{4}{7}+M$	368/7	
x_1 enters	x_2	1/7	1	0	3/7	$-1/7$	22/7	
x_3 leaves	x_3	5/7	0	1	1/7	2/7	26/7	
3 (optimal)	z	0	0	3/5	29/5	$-\frac{2}{5}+M$	$54\frac{4}{5} = z^*$	
	x_2	0	1	$-1/5$	2/5	$-1/5$	$12/5 = x_2^*$	
	x_1	1	0	7/5	$-1/5$	2/5	$26/5 = x_1^*$	

$x_3^* = 0$

$x_4^* = 0$

$R = 0$

Starting Primal Variables	x_4	R
Optimal z-equation coefficient	29/5	$-2/5+M$
Left minus right sides of the dual constraint associated with the starting primal variable	y_1-0	$y_2-(-M)$

Application of the foregoing equation thus yields

$$29/5 = y_1 - 0 \qquad \text{and} \qquad -2/5 + M = y_2 - (-M)$$

We thus obtain $y_1 = 29/5$ and $y_2 = -2/5$, which are the same values as those obtained directly from the optimal dual tableau.

We now show that the optimal *dual* tableau (Table 4-4) also yields the optimal primal solution directly and by using the equation stated earlier. Observe,

* You can reduce the number of variables.
* You can consider it Primal and take the dual, it will return the original problem.

Table 4–4

Iteration	Basic	y_1	y_2'	y_2''	y_3	y_4	y_5	R_1	R_2	R_3	Solution
0 (starting)	w	$-10+4M$	$-8+4M$	$8-4M$	$-M$	$-M$	$-M$	0	0	0	$21M$
	R_1	1	2	−2	−1	0	0	1	0	0	5
	R_2	2	−1	1	0	−1	0	0	1	0	12
	R_3	1	3	−3	0	0	−1	0	0	1	4
4 (optimal)	w	0	0	0	−26/5	−12/5	0	$26/5-M$	$12/5-M$	$-M$	$54\frac{4}{5}$
	y_5	0	0	0	−7/5	1/5	1	7/5	−1/5	−1	3/5
	y_2''	0	−1	1	2/5	−1/5	0	−2/5	1/5	0	2/5
	y_1	1	0	0	−1/5	−2/5	0	1/5	2/5	0	29/5

Return to the original Primal y_1 y_2'' y_2' R_1 R_2 R_3 S_1 S_2 S_3 S_4

$n = 10$

$m = 4$

M cancel in the solution.

$y_2 = y_2' - y_2''$
$= 0 - \frac{2}{5} = -\frac{2}{5}$

first, that x_1, x_2, and x_3 are respectively associated with the first, second, and third *dual* constraints and, hence, with the artificials R_1, R_2, and R_3.†

Starting Dual Variables	$y_1 \leftarrow R_1$	$y_2 \leftarrow R_2$	$y_3 \leftarrow R_3$
Optimal w-equation coefficients	$26/5 - M$	$12/5 - M$	$-M$
Left minus right sides of the primal constraint associated with the starting dual variable	$x_1 - M$	$x_2 - M$	$x_3 - M$

Thus we get

$$x_1 = (26/5 - M) + M = 26/5$$
$$x_2 = (12/5 - M) + M = 12/5$$
$$x_3 = (-M) + M = 0$$

which is the same solution obtained directly in the optimal primal tableau (Table 4–3).

米 Why should we be interested in obtaining the optimal solution of the primal by solving the dual? The answer is that it may be more advantageous computationally to solve the dual rather than the primal. Recall that the computational effort in linear programming depends more on the number of constraints than on the number of variables. Thus, if the dual happens to have a smaller number of constraints than the primal, generally it will be more efficient to solve the dual, from which the optimal primal solution can then be obtained.

Exercise 4.2–1

Find the optimal dual solution from the optimal primal tableau of each of the following examples in Chapter 3.

(a) Example 3.3–1, Table 3–2.
[*Ans.* $y_1 = y_2 = 3/2$.]

(b) Example 3.3–2, Table 3–3.
[*Ans.* $y_1 = 5/8$, $y_2 = 1/8$, $y_3 = 0$.]

(c) The example in Table 3–1, Section 3.2.3–A.
[*Ans.* $y_1 = 7/5$, $y_2 = 0$, $y_3 = -1/5$.]

The primal and dual solutions in Tables 4–3 and 4–4 reveal two interesting results:

1. At the optimum iteration we have

$$\max z = \min w = 54\tfrac{4}{5}$$

This is always true and, indeed, should be consistent with the *optimal* values of the variables in both problems, namely,

$$z = 5x_1 + 12x_2 + 4x_3 = 5 \times 26/5 + 12 \times 12/5 + 4 \times 0 = 54\tfrac{5}{4}$$
$$w = 10y_1 + 8y_2 = 10 \times 29/5 + 8 \times (-2/5) = 54\tfrac{5}{4}$$

† This statement is based on the (almost) obvious observation that *the dual of the dual is the primal.* You should verify this by considering the dual as the *given* problem. Then use x_1, x_2, x_3 as its "dual" variables. By applying the rules in Table 4–2, you will find that the resulting "dual" is the original primal!

H.w Take tablue for slack and artificial

Constrained non linear optimization problems (LNP) only Min.	Monday 8/4/2019

Min $P(\underline{x})$ $\quad \underline{x} = [x_1, x_2 \cdots x_n]^T$

If the feasible region is point there is no optimization

s.t:

$\underline{F}(\underline{x}) = \underline{0}$

$\underline{g}(\underline{x}) \leq \underline{0}$

$\underline{x}_{min} \leq \underline{x} \leq \underline{x}_{max}$

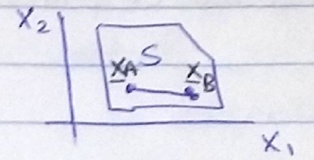
→ Check the convexity of $p(\underline{x})$ using Eigen values of $\underline{\underline{H}}(\underline{x})$ of $p(\underline{x})$

You can classify each optimum due to the eigen values at this point.

* The convexity of the feasible region:

min $p(\underline{x})$

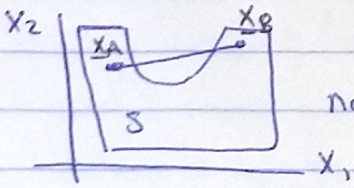
no $\underline{F}(\underline{x})$ s.t: $\underline{g}(\underline{x}) \leq \underline{0}$



S: feasible region convex ✓

convex × ~~the~~ nonconvex

Definition: The feasible region, S, is convex if $\forall_{all}$ $\underline{x}_A$ and $\underline{x}_B \in S$, the joining line between $\underline{x}_A$ and $\underline{x}_B \in S$.



non convex

This means that $\forall$ $\underline{x}_A$ and $\underline{x}_B \in S$

$$\underline{x} = \lambda \underline{x}_A x_A + (1-\lambda)\, \underline{x}_B \in S$$

$$\forall \lambda : 0 \leq \lambda \leq 1$$

Rule: The feasible regions bounded by $\underline{g}(\underline{x})$ where all $\underline{g}(\underline{x})$ are concave functions is a convex region.

strictly concave

$\underline{\underline{H}}$ → Find Eigen values, if +ve or zeros → concave

→ remember that you can use $\underline{\underline{H}}(\underline{x})$ to check the convexity of each function of $\underline{g}(\underline{x})$

Rule: If $p(\underline{x})$ is convex over a convex feasible region, the ~~global~~ local minima is also global minima.

if all λ eigen values in $\underline{\underline{H}}$ are zeros, you can consider it $\begin{cases}\text{concave} \\ \text{convex}\end{cases}$

→ The linear programming problems LPP are convex optimization problem.

if $\underline{\underline{H}} \times -1 \rightarrow$ eigen values don't change.

Example (in book ch.4) * 4.8 example

does the following set of constrains that form the closed feasible region convex?

$-x_1^2 + x_2 \geq 1$ — writing in standard form is better — $g_1(\underline{x}) = -x_1^2 + x_2 - 1 \geq 0$

$x_1 - x_2 \geq -2$ — $g_2(\underline{x}) = x_1 - x_2 + 2 \geq 0$

① $\underline{\underline{H}}_1(\underline{x})$ for $g_1(\underline{x}) \Rightarrow \underline{\underline{H}}_1(\underline{x}) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$

$|\underline{\underline{H}}_1(\underline{x}) - \lambda \underline{\underline{I}}| = 0 \qquad (-2-\lambda)(0-\lambda) - 0 = 0 \qquad -\lambda(-2-\lambda) = 0$

$\boxed{\lambda_1 = 0 \quad \lambda_2 = -2}$ −ve definite

$\rightarrow \underline{\underline{H}}_1(\underline{x})$ negative semi definite $\rightarrow g_1(\underline{x})$ is concave

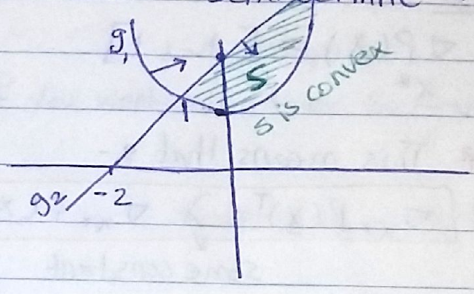
② $\underline{\underline{H}}_2(\underline{x})$ for $g_2(\underline{x}) \Rightarrow \underline{\underline{H}}_2(\underline{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (linear) $\boxed{\lambda_1 = 0 = \lambda_2}$

semi definite < +ve / −ve

اذا اوا ان

$\rightarrow g_2(\underline{x})$ is concave

So, The region is convex.



** Methods of solving non linear programming problems

Ⓐ Direct substitution method — حذف وتعويض

→ For optimization with equality constrains ($\underline{F}(\underline{x})$) ONLY

→ min $p(\underline{x})$ — n decision variables & m equations

s.t : $\underline{F}(\underline{x}) = 0 \qquad \underline{F} = [F_1 \; F_2 \dots F_m]^T$

→ eliminate m variables from $p(\underline{x})$ to convert the problem to unconstrained one with n−m variables

Example min $p(\underline{x}) = 4x_1^2 + 5x_2^2$ — $n = 2 \quad m = 1$ ما عنيى أكثر

s.t : $2x_1 + 3x_2 = 6$

$x_1 = \dfrac{6 - 3x_2}{2} \rightarrow \min p(\underline{x}) = 4\left(\dfrac{6-3x_2}{2}\right)^2 + 5x_2^2 = 14x_2^2 - 36x_2 + 36$

$\boxed{\min p(\underline{x}) = 14x_2^2 - 36x_2 + 36} \qquad \dfrac{\partial P}{\partial x_2} = 0$

$\Rightarrow 28x_2^* - 36 = 0 \qquad \boxed{x_2^* = 36/28 = 1.286 \qquad x_1^* = \dfrac{6 - 3x_2}{2}}$

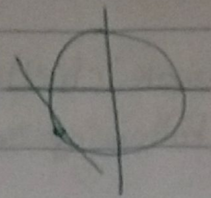
single variable unconstrained

اذا n>2 يكون multivariable

(B) Langrangian method

Example $\min p(\underline{x}) = x_1 + x_2 \nearrow \underline{H} = \underline{0}$ لا يوجد max & min

s.t $F(\underline{x}) = x_1^2 + x_2^2 = 1$



$x_1 = \pm\sqrt{1 - x_2^2}$

<u>caseA</u> $x_1 = -\sqrt{1-x_2^2}$

$\min p(\underline{x}) = -\sqrt{1-x_2^2} + x_2$

$\left.\frac{dp}{dx_2}\right|_{x_2^*} = 0 \qquad -\frac{-2x_2}{2\sqrt{1-x_2^2}} + 1 = 0 \qquad \frac{x_2}{\sqrt{1-x_2^2}} + 1 = 0 \qquad \boxed{x_2^* = \frac{-1}{\sqrt{2}}}$

$\boxed{x_1^* = \frac{-1}{\sqrt{2}}}$ Minima

<u>case B</u> $x_1 = \sqrt{1-x_2^2} \qquad \boxed{x_1^* = x_2^* = \frac{1}{\sqrt{2}}}$ Maxima

<u>Case A</u>

$\nabla_{\underline{x}^*} p(\underline{x}) = [1 \quad 1] \qquad \nabla_{\cancel{x^*}} f(\underline{x}) = [2x_1 \quad 2x_2]^T$

$\nabla_{x^*} f(\underline{x}) = [-\sqrt{2} \quad -\sqrt{2}]^T = -[1.414 \quad 1.414]^T$

$= -1.414 \; [1 \quad 1]^T$

* This means that :-

$\boxed{\nabla_{x^*} P(\underline{x})^T = \lambda \; \nabla_{x^*} f(\underline{x})}$ or F^T $\qquad \lambda = \frac{-1}{1.414}$ minima

some constant

here λ is Langrangian multiplier

$\nabla_{x^*} p(\underline{x})$ and $\nabla_{x^*} f(\underline{x})$ are co-linear vectors

نفس المقدار بس متعاكسين

<u>Case B</u> $\lambda = \frac{1}{1.414}$ maxima

wrong ✗ <u>Note</u> +ve λ corresponds to maxima (even min. problem), while -ve corresponds to minima

$\nabla^T_{\underline{x}^*} P(\underline{x}) - \lambda \nabla_{x^*} F(\underline{x}) = 0$

this helps to build a Langranglan Function $L(\underline{x})$

$\boxed{L(\underline{x}, \lambda) = P(\underline{x}) - \lambda f(\underline{x})}$

$\min p(\underline{x}) \quad \text{s.t}: f(\underline{x})$

To convert to uncons. $L(\underline{x}) = p(\underline{x}) - \lambda f(\underline{x}) = 0$

$\Rightarrow$ Necessary conditions

$\nabla_{\underline{x}^*} p(\underline{x}) = 0^T \qquad F(\underline{x}^*) = 0 \qquad \nabla_{\underline{x}^*} L(\underline{x}) = \underline{0} = \nabla_{x^*} p(\underline{x})$

since $\nabla_{x^*} F(\underline{x}) = 0$

$\nabla_{\underline{x}^*, \lambda} L(\underline{x}) = 0$

Back to example

$\min\ P(\underline{x}) = x_1 + x_2$

s.t $x_1^2 + x_2^2 = 1 \Rightarrow F(\underline{x}) = x_1^2 + x_2^2 - 1 = 0$

$L(\underline{x}, \lambda) = x_1 + x_2 - \lambda(x_1^2 + x_2^2 - 1)$

Sometimes hard system of equations (maybe nonlinear)

$\frac{\partial L}{\partial x_1} = 0 \rightarrow 1 - 2\lambda^* x_1^* = 0$ ①

solve 3 equations by 3 unknown

$x_1^* = 1/2\lambda^*$

$\frac{\partial L}{\partial x_2} = 0 \rightarrow 1 - 2\lambda^* x_2^* = 0$ ②

$x_2^* = 1/2\lambda^*$

$\frac{\partial L}{\partial \lambda} = 0 \rightarrow x_1^{*2} + x_2^{*2} - 1 = 0$ ③

$\left(\frac{1}{2\lambda^*}\right)^2 + \left(\frac{1}{2\lambda^*}\right)^2 - 1 = 0$

$\lambda^2 = \frac{1}{2}$ $\boxed{\lambda = \pm \frac{1}{\sqrt{2}}}$

هي التي تحدد الحل

if you want min take the -ve sign λ

Wednesday 10/4/2019

$\boxed{\begin{array}{l} \min\ P(\underline{x}) = \cdots \\ \text{s.t } F(\underline{x}) = 0 \end{array}}$

constrained to unconstrained

$L(\underline{x}, \lambda) = P(\underline{x}) - \lambda F(\underline{x})$

at $\underline{x} = \underline{x}^*$ $F(\underline{x}^*) = 0$

$L(\underline{x}^*, \lambda^*) = P(\underline{x}^*) - \lambda^* \cancel{F(\underline{x}^*)}^{\text{zero}}$

$\nabla_{\underline{x}, \lambda} L(\underline{x}, \lambda) = \underline{0}^T$

* In this course, the gradient of scalar is row vector

$\nabla_{\underline{x}, \lambda} L(\underline{x}, \lambda) = \lambda \nabla P(\underline{x})$

موجودة max. موجودة min.

λ : Langrangian multiplier which can take +ve or -ve values

Example

You can make it graphically

min. $P(\underline{x}) = 4x_1^2 + 5x_2^2 \rightarrow$ ellipse $\Rightarrow \frac{P}{20} = \frac{x_1^2}{5} + \frac{x_2^2}{4}$

s.t $\quad 2x_1 + 3x_2 = 6 \rightarrow$ -ve slope

$F(\underline{x}) = 2x_1 + 3x_2 - 6 = 0$ (to make $L(x^*, \lambda^*) = P(\underline{x}^*)$)

non linear function & linear constrains.

$L(\underline{x}, \lambda) = 4x_1^2 + 5x_2^2 - \lambda(2x_1 + 3x_2 - 6)$

f نواحي

$\nabla_{\underline{x}, \lambda} L = 0^T$

$\frac{\partial L}{\partial x_1} = 0 = 8x_1 - 2\lambda \rightarrow (1)$

$\frac{\partial L}{\partial x_2} = 0 = 10x_2 - 3\lambda \rightarrow (2)$

نفس $F(\underline{x})$ $\quad \frac{\partial L}{\partial \lambda} = 0 = 2x_1 + 3x_2 - 6 \rightarrow (3)$

3 equations & 3 unknowns
linear system

$x_1^* = 1.017$

$x_2^* = 1.286$

$\boxed{\lambda^* = 4.286}$ +ve λ But min !

- if the problem is (a) max $\Rightarrow$ it will be at ∞, multi-solutions $\frac{1}{6}(2)$ but both goes to ∞

* Langrangian method for $P(\underline{x})$ and with more than 1 equality constrains

min $P(\underline{x})$

s.t : $\underline{F}(\underline{x}) = \underline{0} \rightarrow \underline{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

$L(\underline{x}, \underline{\lambda}) = P(\underline{x}) - \underbrace{\underline{\lambda}^T}_{1\times m} \underbrace{\underline{F}(\underline{x})}_{m\times 1}$

λ_1 for F_1 $\quad$ • λ could be +ve or negative.

λ_2 for F_2

$\vdots$

λ_m for F_m

→ Necessary conditions

$\nabla_{\underline{x}, \underline{\lambda}} L(\underline{x}, \underline{\lambda}) = 0^T$

Example $\min\ p(\underline{X}) = -X_1^2 + X_2^2 + X_3^2$ $n = 3$

s.t: $X_1 + 2X_2 + 3X_3 = 1$ $m = 2$

$X_1^2 + \frac{X_2^2}{2} + \frac{X_3^2}{4} = 4$

You can solve it in direct substitution method to make it single variable, but problem is what variable we must eliminate it differs ? Because it is non linear. يكون مختلف

$L(\underline{X}, \underline{\lambda}) = -X_1^2 + X_2^2 + X_3^2 + -\lambda_1(X_1 + 2X_2 + 3X_3 - 1)$
$- \lambda_2(X_1^2 + \frac{1}{2}X_2^2 + \frac{1}{4}X_3^2 - 4)$

$\frac{\partial L}{\partial X_1} = 0 = -2X_1 - \lambda - 2\lambda_2 X_1 \longrightarrow ①$

$\frac{\partial L}{\partial X_2} = 0 = 2X_2 - 2\lambda_1 - \lambda_2 X_2 \longrightarrow ②$

$\frac{\partial L}{\partial X_3} = 0 = 2X_3 - 3\lambda_1 - \frac{1}{2}\lambda_2 X_3 \longrightarrow ③$

$\frac{\partial L}{\partial \lambda_1} = 0 = X_1 + 2X_2 + 3X_3 - 1 \longrightarrow ④$

$\frac{\partial L}{\partial \lambda_2} = 0 = X_1^2 + \frac{X_2^2}{2} + \frac{X_3^2}{4} - 4 \longrightarrow ⑤$

- 5 NL equations by 5 unknowns, on Polymath or least square method solve to get $\underline{X}^*$ and $\underline{\lambda}^*$
  if more than 1 solution, substitute in $p(X)$ and take the minimum

(سؤال الـ global)

if $\lambda_1^* = 200$ $\lambda_2^* = 1.5$ means that the ~~problem~~ optima is sensitive to λ_1^* (RHS constants) $\Rightarrow$ related to assumptions if safe or not

usually RHS is related to distribution.

Note

$\min\ P(\underline{X})$

S.t $\underline{F}(\underline{X}) \mp \underline{e} = 0$ $\underline{e} = [e_1\ e_2\ \cdots\ e_m]^T$

$L(\underline{X}, \underline{\lambda}) = P(\underline{X}, \lambda) - \underline{\lambda}^T[\underline{F}(\underline{X}) \mp e]$

$\nabla_{\underline{e}} L(\underline{X}, \lambda) = 0 - -\underline{\lambda}^T = \underline{\lambda}^T$

$\frac{\partial L}{\partial e_1} = \lambda_1$; $\frac{\partial L}{\partial e_2} = \lambda_2$ .... $\frac{\partial L}{\partial e_m} = \lambda_m$

@ $\underline{X} = \underline{X}^* \Rightarrow p(\underline{X}^*) = L(\underline{X}^*) \Rightarrow \frac{\partial P}{\partial e_1} = \frac{\partial L}{\partial e_1} = \lambda_1$

$\frac{\Delta P}{\Delta e_1} = \lambda_1$ $\Delta P = \Delta e_1 \lambda_1 = P^{new} - P^*$

if sensitive, big change relatively

* Langrange method for equality and inequality

$\min \ p(\underline{x})$

$\text{s.t.} \quad \underline{F}(\underline{x}) = \underline{0}$

$\underline{g}(\underline{x}) \geq \underline{0}$ ← always for $\geq$, if $\leq$ convert

$L(\underline{x}, \underline{\lambda}, \underline{\mu}) = p(\underline{x}) - \underline{\lambda}^T \underline{F}(\underline{x}) - \underline{\mu}^T \underline{g}(\underline{x})$

$\underline{\lambda}$ Langrangian multipliers for $\underline{F}(\underline{x})$, they can be +ve or −ve

$\underline{\mu}$ Langrangian multipliers for $\underline{g}(\underline{x})$, they must be greater than or equal to zero
For minimize

Necessary conditions

* $\nabla_{\underline{x}, \underline{\lambda}, \underline{\mu}} L(\underline{x}, \underline{\lambda}, \underline{\mu}) = 0^T$

* P, $\underline{F}$, and $\underline{g}$ are twice differentiable at $\underline{x}^*$ (stationarity)

In Final Exam

→ Karash – Kuhn – Tucker (KKT) condition

لو تأكد من الحل يعوضها للي $\underline{x}^*$

① stationarity → of $p(x)$

$$\nabla_{\underline{x}} L \Big|_{\underline{x}^*, \underline{\lambda}, \underline{\mu}} = \underline{0}^T$$

② Feasibility : The optimum solution must satisfy both $\underline{F}(\underline{x})$ ang $\underline{g}(\underline{x})$ constrains $\nabla_{\underline{\lambda}\,\underline{\mu}} L(\underline{x}, \underline{\lambda}, \underline{\mu}) = \underline{0}^T$

③ Constrains qualifications : for $\underline{g}(\underline{x}) \left\langle \begin{array}{l} g_I \text{ (inactive)} \\ g_A \text{ (active)} \end{array} \right.$

The equality constrains $\underline{F}(\underline{x})$, and the active inequality ($\underline{g}(\underline{x}) = 0$) $g_A(\underline{x})$

عدد المتغيرات = عدد المعادلات من F و g_A

are linearly independent . (μ is usefull here)

$g_A(x) = \underline{0}$, $\underline{F}(\underline{x}) = \underline{0}$ → must be independent

$g_I(x) > \underline{0}$

④ Non-negativity $\underline{\mu} \geq 0$

⑤ Complimentarity

ما بروح الاثنين = صفر

$\mu_i g_i(\underline{x}^*) = 0 \quad : \quad \left[\begin{array}{ll} \mu_i^* = 0 & g_i(\underline{x}^*) > 0 \\ \mu_i^* > 0 & g_i(\underline{x}^*) = 0 \rightarrow g_A(\underline{x}) \end{array} \right.$

⑥ Curvature

The $\underline{H}(x)$ of $\underline{L}(\underline{x}, \underline{\lambda}, \underline{\mu})$ is +ve semi-definite if :

$$\underline{v}^T \nabla^2_{\underline{x}\underline{x}} L(x, \lambda, \mu) \, \underline{v} \Big|_{\underline{x}^*, \underline{\mu}^*, \underline{\lambda}^*} \geq 0 \quad \text{where}$$

$$\underbrace{\begin{bmatrix} \nabla_{\underline{x}} F(\underline{x}) \\ \\ \nabla_x g_A(x) \end{bmatrix}_{n\times n}}_{\text{Jacobian}} \underset{n\times 1}{\underline{V}} = \underline{0}$$

$$\begin{bmatrix} \nabla_{\underline{x}} F(\underline{x}) \\ \\ \nabla_x g_A(x) \end{bmatrix}_{n \times n} \underline{V}_{n \times 1} = \underline{0}$$

Jacobian

Monday 15/4

Example : $P(\underline{x}) = (x_1 - 1)^2 + x_2^2$ it is open region

s.t : $x_1 - x_2^2 \le 0 \rightarrow g_1(\underline{x}) = x_2^2 - x_1 \ge 0$

$L(\underline{x}, \underline{\lambda}, \underline{\mu}) = p(\underline{x}) - \underline{\lambda}^T \underbrace{F(\underline{x})}_{\text{zero}} - \underline{\mu}^T g(\underline{x})$

$L(\underline{x}, \mu) = (x_1 - 1)^2 + x_2^2 - \mu(x_2^2 - x_1)$

$\left.\frac{\partial L}{\partial x_1}\right| = 0 \qquad 2(x_1 - 1) + \mu = 0 \; \ldots\ldots \; (1)$

$\frac{\partial L}{\partial x_2} = 0 \qquad 2x_2 - 2\mu x_2 = 0 \; \ldots\ldots \; (2)$

$\frac{\partial L}{\partial \mu} = 0 \qquad -2x_2 + x_1 = 0 \; \ldots\ldots \; (3)$

solve the above non-linear three equations to get the three solutions :-

(1) $x_1^* = 1/2 \quad x_2^* = \sqrt{1/2} \quad \mu^* = 1$

(2) $x_1^* = 1/2 \quad x_2^* = -\sqrt{\frac{1}{2}} \quad \mu^* = 1$

(3) $x_1^* = 0 \quad x_2^* = 0 \quad \mu^* = 2$

$P^*_{(1)} = 0.75 \qquad P^*_{(2)} = 0.75 \qquad P^*_{(3)} = 1$

which is maxima and which is minima ?

global or local ? $\Rightarrow$ by curvature checking condition

since $\uparrow \mu$, all are non zero

✓ stationarity ✓ visibility constrains qualifications ✓ complimintary ✓

• check curvature

$\underline{V}^T \nabla^2_{\underline{x},\underline{x}} L(\underline{x}, \underline{\mu}, \underline{\lambda})\Big|_{x^*, \mu^*, \lambda^*} \underline{V} \ge 0$ L is +ve semi-definite $\downarrow$ local minima

if > it is global minima .

if no f $\Rightarrow$ zero

أولوية أكثر لـ g

where $\begin{bmatrix} \nabla \underline{F}(x) \\ \nabla \underline{g}(x) \end{bmatrix} \underline{V} = \underline{0}$

$\rightarrow$ the third solution $\quad x_1^* = x_2^* = 0 \quad \mu^* = 2 \quad \underline{V}^T = [V_1 \;\; V_2]$

$\nabla^2_{\underline{x},\underline{x}} L(\underline{x}, \underline{\mu}, \underline{\lambda})\Big|_{\underline{x}^*, \underline{\mu}^*} = \begin{bmatrix} 2 & 0 \\ 0 & 2-2\mu \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

$\nabla \underline{g}(\underline{x})\Big|_{x^*} = [-1 \quad 2x_2] = [-1 \quad 0]$

$\begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$V_1 \times 0 + 0 \times V_2 = 0$

$-V_1 + 0 V_2 = 0 \quad \boxed{V_1 = 0} \quad V = \begin{bmatrix} 0 \\ V_2 \end{bmatrix}$

$$[0 \quad V_2]\begin{bmatrix}2 & 0\\ 0 & -2\end{bmatrix}\begin{bmatrix}0\\ V_2\end{bmatrix} = [0 \quad -2V_2]\begin{bmatrix}0\\ V_2\end{bmatrix} = 0 - 2V_2^2 = -2V_2^2$$

$$-2V_2^2 \le 0$$

$\Rightarrow$ The third solution is gives the local maxima.

$\rightarrow$ The ~~second~~ First solution $x_1^* = 1/2 \quad x_2^* = -\sqrt{1/2} \quad \mu^* = 1$

$\underline{V}^T \ \nabla^2 L \ \underline{V} = 4V_2^2 \geq 0$

$\Rightarrow$ the first solution gives local minima.

© Penalty Function method (only for min.)

$\rightarrow$ You convert the constrained optimization problem to unconstrained one using some penalty function

Min $P(\underline{x})$

s.t : $\underline{F}(\underline{x}) = \underline{0}$

$\underline{g}(\underline{x}) \geq 0$

$$\boxed{\text{Min } \phi(\underline{x}, r) = P(\underline{x}) + \beta(\underline{F}(\underline{x}), \underline{g}(\underline{x}), r)}$$

$\longrightarrow$ where r is penalty parameter must be +ve since min.

You search to find r, $r \geq 0$ Big M technique زي

• Common penalty function used :-

Ⓐ Quadratic penalty function

$$\text{Min } \phi(\underline{x}, r) = P(\underline{x}) + \frac{r}{2}\underbrace{\left[\sum_{i=1} F_i^2(\underline{x}) + \sum_{i=1}[\min(0, g_i(\underline{x}))]^2\right]}_{\text{must be zero at optimum}}$$

because of derivation

@ $\underline{x} = \underline{x}^* \Rightarrow \phi^* = P^*$

Ⓑ Absolute penalty function

$$\text{Min } \phi(\underline{x}, r) = \underline{P}(\underline{x}) + r\left[\sum_{i=1} |F_i(\underline{x})| + \sum |\min(0, g_i(\underline{x}))|\right]$$

Example $\min P(\underline{x}) = (x_1 - 1)^2 + (x_2 - \cancel{4}2)^2$

s.t ~~$f(x) =$~~ $x_1 + x_2 = 4$ By Quadratic

$F(\underline{x}) = x_1 + x_2 - 4 = 0$

$$\phi(\underline{x}, r) = (x_1 - 1)^2 + (x_2 - \cancel{4}2)^2 + \frac{r}{2}[x_1 + x_2 - 4]^2$$

$$\frac{\partial \phi}{\partial x_1} = 0 = 2(x_1 - 1) + \frac{2r}{2}(x_1 + x_2 - 4)$$

$$\frac{\partial \phi}{\partial x_2} = 0 = 2(x_2 - \cancel{4}2) + \frac{2r}{2}(x_1 + x_2 - 4)$$

You can take $r = 0$ but check the constrains
You can't since you will neglect all the constrains.

r	x_1	x_2	ϕ
0	1	2	0
2	1.3333	2.3333	0.3333
20	1.4762	2.4762	0.4762
200	~~1.4998~~ 1.4975	2.4975	0.4975
2000	1.4998	2.4998	0.4998
↓	1.5 = x_1^*	2.5 = x_2^*	0.5 = $P(x^*)$ = ϕ^*

stop until having specific error when substituting

You can use the curvature to help you

if You put big small r, terms will be like neglected, as you cancle constrains or the objective ~~the~~ function

↓

Drawbacks : ILL - conditioning numerical problem wher r is very large Also, slow convergence

(D) Logarithmic barrier function method

→ for inequality constrained problem

Min $P(\underline{x})$ s.t $g(\underline{x}) \geq 0$ $\Longrightarrow$ Min $B(\underline{x}) = P(\underline{x}) - \rho \sum_{i=1} \ln [g_i(\underline{x})]$

because there is no -ve in ln

ln fraction is -ve

only for active

when being close, $*10^{-7} \Rightarrow \ln \to$ big

when $\underline{x} \to \underline{x}^*$: $\rho \to 0$: $P^* = B^*$

Example min $P(\underline{x}) = (x_1-1)^2 + (x_2-2)^2$

s.t $x_1 + x_2 \geq 4$ $\to$ $g(\underline{x}) = x_1 + x_2 - 4 \geq 0$

$B(\underline{x}, \rho) = (x_1-1)^2 + (x_2-2)^2 - \rho \ln [x_1 + x_2 - 4]$

Derive w.r.t x_1 & x_2

error $= \left| \frac{B-P}{P} \right|$

ρ	x_1	x_2	$P(\underline{x})$	$g(\underline{x})$	$\rho \ln g$	$B(\underline{x}, \rho)$
10	2.851	3.851	6.851	2.702	9.938	-3.088
1	1.809	2.809	1.309	0.618	-0.481	0.556
0.01	1.505	2.505	0.510	0.01	-0.046	0.556
	1.5 = x_1^*	2.5 = x_2^*	0.5	0	0	0.5

I don't want very small ρ because of ill conditioning

(D) Mixed penalty-barrier functions method

min $P(\underline{x})$

s.t $\underline{f}(\underline{x})=0$

$\underline{g}(\underline{x})\geq 0$

$$\phi(\underline{x},r,\rho)=P(\underline{x})+\frac{r}{2}\sum_{i=1}[f_i(\underline{x})]^2-\rho\sum_{i=1}\ln g_i(\underline{x})$$

	$r=0.1$	$r=1$	$r=10$ - - - - -
$\rho=10$			
$\rho=1$			
⋮			

- when you substitute in $f(\underline{x})$ you must $=0$
- You fixed in r, if it diverge, stop and search for ρ.

(E) Successive linear programming (SLP) method

linearization by Taylor series

A. Linearize $P(\underline{x})$ $\underline{f}(x)$ $\underline{g}(x)$ using Taylor series expansion around initial guess of $\underline{x}^*$ — Initial guess $\tilde{\underline{x}}$

- $\tilde{\underline{x}}$ must be within the feasible region → it must satisfy the $\underline{f}(x)$ and $\underline{g}(\underline{x})$ constrains
- min. $P(\underline{x})\cong P(\tilde{\underline{x}})+\nabla_{\underline{x}}P(\underline{x})\big|_{\tilde{\underline{x}}}(\underline{x}-\tilde{\underline{x}})$ ← it is linear

You can remove it and make new function

s.t.:-

$$\underline{f}(\underline{x})\cong\underline{f}(\tilde{\underline{x}})+\nabla_x\underline{f}(\underline{x})\big|_{\tilde{\underline{x}}}(\underline{x}-\tilde{\underline{x}})=0$$

$$\underline{g}(\underline{x})\cong g(\tilde{\underline{x}})+\nabla_{\underline{x}}\underline{g}(\underline{x})\big|_{\tilde{\underline{x}}}(\underline{x}-\tilde{\underline{x}})\geq 0$$

- Limit bounds $-\underline{\Delta}\leq\underline{x}-\tilde{\underline{x}}\leq\underline{\Delta}$ where $\underline{\Delta}$ limits the change in the variable of iterations, they called step bounds. They imposed to ensure that the errors between the non linear and approximate linear one are not too large.

B. Solve the approximate linear programming problem iteratively to get $\underline{x}^*$.

$\Rightarrow$ let $\underline{d}=\underline{x}-\tilde{\underline{x}}=\begin{bmatrix}x_1-\tilde{x}_1\\x_2-\tilde{x}_2\\x_3-\tilde{x}_3\end{bmatrix}$

$\phi(\underline{x})=P(\underline{x})-P(\tilde{\underline{x}})$

min $\phi(\underline{x})=\underline{c}^T\underline{d}$ where $\underline{c}^T=\nabla P(\underline{x})\big|_{\tilde{\underline{x}}}$

subject to :-

$\underline{\underline{A}}^{eq}\, \underline{d} = \underline{b}^{eq}$ where $\underline{\underline{A}}^{eq} = \nabla_{\underline{x}} \underline{F}(\underline{x})|_{\tilde{\underline{x}}}$; $\underline{b}^{eq} = -F(\tilde{\underline{x}})$

$\underline{\underline{A}}^{ineq.}\, \underline{d} \geq \underline{b}^{ineq.}$ where $\underline{\underline{A}}^{ineq.} = \nabla_{\underline{x}}\, \underline{g}(\underline{x})|_{\tilde{\underline{x}}}$; $\underline{b}^{ineq.} = -\underline{g}(\tilde{\underline{x}})$

or $\leq$ no problem

Solve in simplex or Lindo

$\underline{d} \leq \underline{\Delta}$ when you put in Lindo, d must be $d = d' - d''$

$\underline{d} \geq -\underline{\Delta}$

Put error percentage ⇒ which is $\underline{d}$ but absolute value as a ~~test~~

$\sum (error)_i^2 \leq e$

Example Max $2x_1 + x_2 = P(\underline{x})$

s.t. $x_1^2 + x_2^2 \leq 25$

$x_1^2 - x_2^2 \leq 7$

$x_1 \geq 0$

From question $x_2 \geq 0$

↳ Take $\tilde{\underline{x}} = [2 \quad 2]^T$ as an initial guess and $\underline{\Delta} = [1 \quad 1]^T$

$P(\tilde{\underline{x}}) = p(2,2) = 6$

$\phi = P(\underline{x}) - 6$

$\underline{c}^T = \nabla P(\underline{x}) = [2 \quad 1]$

$\underline{d} = [d_1 \quad d_2]^T$

$$\underline{\underline{A}}^{ineq.} = \nabla_{\underline{x}}\, g(\underline{x})|_{\tilde{\underline{x}}} = \begin{bmatrix} 2\tilde{x}_1 & 2\tilde{x}_2 \\ 2\tilde{x}_1 & -2\tilde{x}_2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & -4 \end{bmatrix}$$

$$\underline{b}^{ineq.} = \begin{bmatrix} 25 \\ 7 \end{bmatrix}$$

• x_1^2 $L(x_1) = x_1^2$ but requires $\tilde{x}_1$

Iteration #1 ~~$\phi(\underline{x}) = P(\underline{x}) - 6$~~

max. $\phi(\underline{x}) = 2d_1 + d_2$ where $\underline{c}^T = \nabla P(\underline{x})|_{\tilde{\underline{x}}}$

s.t: ~~$4d_1 + 4d_2$~~

$4d_1 + 4d_2 \leq 25$ You can divide by 4

$4d_1 - 4d_2 \leq 7$ You can divide by 4

$d_1 \leq 1$ $\quad d_1 \geq -1$

$d_2 \leq 1$ $\quad d_2 \geq 1$

solve this LP problem using simplex method to get $\underline{d} = [1 \quad 1]^T$

$\underline{d} = \underline{x} - \underline{\tilde{x}}$ $\quad\quad$ $\underline{x} = \underline{d} + \underline{\tilde{x}}$

$$\underline{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Iteration #2

$\underline{\tilde{x}} = \underline{x}^{(1)} = [3 \quad 3]^T$

to have max. $2d_1 + d_2$

s.t. $6d_1 + 6d_2 \leq 25$

$d_1 - d_2 \leq 7/6$

$d_1 \leq 1$, $d_1 \geq -1$

$d_2 \leq 1$, $d_2 \geq -1$

stop when if $\sum_i^n \frac{|(\underline{x}^{(i+1)} - \underline{x}^{(i)})|}{|\underline{x}^{(i)}|} < \epsilon$ stop $\quad \epsilon = 0.001$ "Percent relative error"

continue

Ⓔ Quadratic Programming (QP) $\quad$ matlab command also

min $P(\underline{x}) = \underline{c}^T \underline{x} + \frac{1}{2} \underline{x}^T \underline{\underline{A}} \underline{x}$ $\quad$ · helpful, since least square method is quadratic

s.t.: $\underline{\underline{A}}^{eq.} \underline{x} = \underline{b}^{eq.}$

$\underline{\underline{A}}^{ineq.} \underline{x} \geq \underline{b}^{ineq.}$ $\quad$ linear constrains

$\underline{x} \geq 0$

matlab command: Quadprog (see its syntax)

in successive linear → convert to quadratic (if complex)
$\underline{\underline{A}}$ terms are found, then QP instead of simplex, and make linearisation of constrains