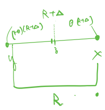
- **optimization problem**

$$\underset{x}{\text{minimize}}\ f_0(x) \quad \text{s.t.}\ f_i(x) \le 0,\ h_i(x) = 0$$

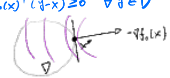
- feasible: $x \in \mathrm{dom}\, f_0$
- infeasible: $p^* = \infty / -\infty$
- optimal value
  $p^* = \min\{f_0(x) \mid f_i(x) \le 0,\ h_i(x) = 0\}$
  x^* is optimal
- locally optima
  $f_0(x) = \inf\{f_0(z) \mid z \in D,\ \|z-x\|_2 \le R\}$
  $D = \bigcap_{i=0}^{m} \mathrm{dom} f_i \cap \bigcap_{i=0} \mathrm{dom}\, h_i$
- feasibility problem
  find x s.t. $f_i(x) \le 0$, $h_i(x) = 0$ $\iff$ $\underset{x}{\text{minimize}}\ 0$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$
- convex optimization
  $\underset{x}{\text{minimize}}\ f_0(x)$ → convex
  s.t. $f_i(x) \le 0$ → convex
  $a_i^T x = b_i$ → affine
- quasiconvex optimization
  $\underset{x}{\text{minimize}}\ f_0(x)$ → quasiconvex
  s.t. $f_i(x) \le 0$, $a_i^T x = b_i$
  $S_\alpha = \{x \in D \mid f(x) \le \alpha\}$
  $\alpha = R$
  S_α = convex
  $f(x) \in$ quasiconvex
- local & global optima
  $\underset{z}{\text{minimize}}\ f_0(z)$
  s.t. $f_i(z) \le 0$, $h_i(z) = 0$
  $\|z - x\|_2 \le R$
  x is locally optimal when
  $f_0(x) = \inf\{f_0(z) \mid z \text{ feasible}, \|z-x\|_2 \le R\}$ ← fall back here

in convex
if x is local optimal
then x is global optimal
proof:
△ if $\exists y$: $f_0(y) < f_0(x)$, $\|y - x\|_2 > R$ } as if $\|y-x\|_2 < R$
△ we let $\theta = \dfrac{R}{2\|y-x\|_2} < \dfrac{1}{2}$
△ we let
$z = (1-\theta)x + \theta y$
convex combination, in convex function
⇒ ∴
$f_0(z) \le (1-\theta) f_0(x) + \theta f_0(y)$
$0 < \theta < \frac{1}{2}$



⇒ from here
$\|x - z\|_2 = \theta(R + \Delta)$
$\le \frac{1}{2}(R + \Delta)$
• we let $\|x - z\|_2 = \frac{1}{2}R$ to fullfil all the condition (we want z to be within $\|x-z\| < R$, s.t. x is local optimal)
△ w/ assumption $f_0(y) < f_0(x)$
we also hv
$\theta f_0(y) + (1-\theta) f_0(x) < f_0(x)$
⇒ $f_0(z) = \theta f_0(y) + (1-\theta) f_0(x) < f_0(x)$
contradicts w/
$f_0(x) = \inf\{f_0(z) \mid z \in D, \|z - x\|_2 \le R\}$

- optimality criterion
  $\nabla f_0(x)^T (y - x) \ge 0 \quad \forall y \in D$

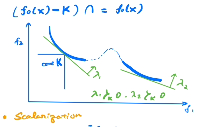


- **Geometric Programming**
  - monomial function
    $f: \mathbb{R}^n \to \mathbb{R}$, $\mathrm{dom} f = \mathbb{R}^n_{++}$
    $f(x) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$
    $c > 0$, $a_i \in \mathbb{R}$
  - posynomial function
    $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$
    $c_k > 0$, $a_{ik} \in \mathbb{R}$
    sum of monomial function
  - transfer to convex problem
    $y_i = \log x_i$, $x_i = e^{y_i}$
    $f(x) = f(e^{y_1}, \dots, e^{y_n})$
    $= c(e^{y_1})^{a_1} \cdots (e^{y_n})^{a_n}$
    $= e^{a^T y + b}$ $(b = \log c)$
    monomial

original:
$\underset{}{\text{minimize}}\ f_0(x)$
s.t. $f_i(x) \le 1$ (posy)
$h_i(x) = 1$ (mono)

convex equivalence:
minimize $\log\left(\sum_{k=1}^{K} e^{a_{0k}^T y + b_{0k}}\right)$
s.t. $\log\left(\sum_{k=1}^{K} e^{a_{ik}^T y + b_{ik}}\right) \le 0$
$Gy + d = 0$

- **multicriterion optimization**
  minimize $f_0(x) = [F_1(x), \dots F_q(x)]$
  s.t. $f_i(x) \le 0$, $Ax = b$
  - Pareto optimal
    $(f_0(x) - K) \cap = f_0(x)$



  - Scalarization
    minimize $\lambda^T f_0(x)$
    s.t. $f_i(x) \le 0$, $h_i(x) = 0$

⇒ $\forall x^{po}$,
$\exists \lambda \succ_{K^*} 0$
s.t. $\tilde{x}$ is an optimal point of the scalarized problem

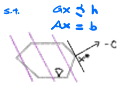
- **Duality**
  - The Lagrangian
    minimize $f_0(x)$ → λ_i
    s.t. $f_i(x) \le 0$, $i = 1, \dots, m$
    $h_i(x) = 0$, $i = 1, \dots, p$ → ν_i
    $L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$
  - The Lagrange Dual Function
    $g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$
    ∴ $g(\lambda, \nu) \le p^* = f_0(\tilde{x})$ → all feasible x
    ($\lambda \succeq 0$; lower bound)
    ∴ $\lambda \succeq 0$, $(\lambda, \nu) \in \mathrm{dom}\, g$
    dual feasible
  - The Lagrange Dual Problem
    maximize $g(\lambda, \nu)$
    s.t. $\lambda \succeq 0$

- Weak Duality
  $d^* \le p^*$
  (best bound, optimal value)
  △ = duality gap
- Strong Duality
  - $d^* = p^*$
  - Slater's condition
    $\exists x \in \mathrm{relint}\, D$
    s.t. $f_i(x) < 0$, $Ax = b$

- optimality conditions – complementary slackness
  assume strong duality holds:
  $d^* = p^*$
  $g(\lambda^*, \nu^*) = f_0(x^*)$
  ⇒
  $f_0(x^*) = g(\lambda^*, \nu^*)$
  $= \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x)\right)$
  $\le f_0(x^*) + \sum_{i=1}^{m} \underbrace{\lambda_i^*}_{\ge 0} \underbrace{f_i(x^*)}_{\le 0} + \sum_{i=1}^{p} \nu_i^* \underbrace{h_i(x^*)}_{=0}$
  $\le f_0(x^*)$
  {if $\lambda_i^* > 0 \Rightarrow f_i(x^*) = 0$
  {if $f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0$
  complementary slackness!
  ∴ $\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$
- KKT optimality conditions
  - let x^* (primal), (λ^*, ν^*) (dual) be optimal points
  - 0-duality gap
  ∴ $x^* = \underset{x}{\arg\min}\, L(x, \lambda^*, \nu^*)$
  - $\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$
  ∴
  - $f_i(x^*) \le 0$
    $h_i(x^*) = 0$
    $\lambda_i^* \ge 0$
    $\lambda_i^* f_i(x^*) = 0$
  → KKT optimality conditions!

- **Sufficiency & Necessity**
  if P then Q
  △ Q is necessary for P
  ① truth of Q is guaranteed by truth of P
  ② (falsity of Q guarantees falsity of P)
  ▲ P is sufficient for P
  ③ truth of P guarantees truth of Q
  ④ but falsity of P might still hv truth of Q
  如果"要選總統"(P)那"要40歲"(Q)
  e.g.
  ① "人一定是40歲"(Q)如果"這人選總統"(P)
  ② 如果"未40歲"(Q)那"一定非總統"(P)
  ③ 如果"要選總統"(P)那"一定40歲"(Q)
  ④ 如果"不選總統"(P)那"可能40歲"(Q)
  sum:
  if P then Q
  △ Q is P's necessary condition
  ▲ P is Q's sufficient condition

- **Convex problems**
  - Linear Programming
    $\underset{x}{\text{minimize}}\ c^T x + d$
    s.t. $Gx \preceq h$, $Ax = b$



  - piecewise-linear minimization
    $f_0(x) = \max_{i=1,\dots,m} (a_i^T x + b_i)$
    $\underset{x}{\text{minimize}}\ f_0(x)$
    ⇕
    minimize t
    s.t. $a_i^T x + b_i \le t$, $i = 1, \dots, m$
  - Quadratic Program
    $\underset{x}{\text{minimize}}\ \frac{1}{2} x^T P x + q^T x + r$
    s.t. $Gx \preceq h$, $Ax = b$
    $P \in S^n_+$
  - Quadratically Constrained Quadratic Program
    $\underset{x}{\text{minimize}}\ \frac{1}{2} x^T P_0 x + q_0^T x + r_0$
    s.t. $\frac{1}{2} x^T P_i x + q_i^T x + r_i \le 0$
    $Ax = b$
    $P_i \in S^n_+$ (over a feasible region that's intersection of ellipsoids + affine set)

- Second-order Cone Programming
  $\underset{x}{\text{minimize}}\ f^T x$
  $\|A_i x + b_i\|_2 \le c_i^T x + d_i$
  $Fx = g$
  SOCP could be understood:
  $\begin{bmatrix} Ax+b \\ c^T x + d \end{bmatrix} \in \mathbb{R}^{k+1}$ lies in SOC
  $(Ax + b, c^T x + d) \succeq_Q 0$
  G.I. for cone Q
- e.g. Robust Linear Programming

- **Perturbation & Sensitivity Analysis**
  - recall
    $\underset{x}{\text{minimize}}\ f_0(x)$
    s.t. $f_i(x) \le 0$, $h_i(x) = 0$
    perturbed by u_i, v_i
    $\underset{x}{\text{minimize}}\ f_0(x)$
    s.t. $f_i(x) \le u_i$, $h_i(x) = v_i$
  - optimal value (becomes a function) w.r.t. (u, v)
    $p^*(u, v) = \inf\{f_0(x) \mid \exists x \in D,\ f_i(x) \le u_i,\ h_i(x) = v_i\}$
  - let (λ^*, ν^*) → dual optimal pt. of unperturbed problem
    ⇒ $p^*(0,0) = g(\lambda^*, \nu^*)$
    $\le f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x)$
    $\le f_0(x) + \lambda^{*T} u + \nu^{*T} v$
    ⇒ $p^*(0,0) \le f_0(x) + \lambda^{*T} u + \nu^{*T} v$
    ⇒ $f_0(x) \ge p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$
    ⇒ $p^*(u,v) \ge p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$
    $(p^*(u,v) \in \{f_0(x) \mid x \in \mathrm{dom} f\})$
    $p^*(u,v) \ge p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$
    perturbation analysis

- local sensitivity
  assume
  - $p^*(u,v)$ differentiable @ $u=0$, $v=0$
  - strong duality holds
  ⇒ $p^*(u,v)$ @ $u=0, v=0$
  $\dfrac{\partial p^*(0,0)}{\partial u_i} = -\lambda_i^*$

  $\dfrac{\partial p^*(0,0)}{\partial v_i} = -\nu_i^*$
- w/ complementary slackness:
  $\lambda_i f_i(x) = 0$
  - $f_i(x) < 0$, $\lambda_i = 0$
    then: value @ $p^*(0,0)$ doesn't change w.r.t. u
  - $f_i(x) = 0$, $\lambda_i^* > 0$
    then: value @ $p^*(0,0)$ changes w.r.t. u
    $\lambda_i^* \uparrow$ → perturb ↑
    $\lambda_i^* \downarrow$ → perturb ↓
  → tells us HOW ACTIVE IT IS

• Numerical Method of Differential Equations

• ODE 4 types

1 separable equations

$\frac{dy}{dx} = P(x)Q(x)$

e.g. $y' + 2xy = x$

$\Rightarrow \frac{dy}{dx} + 2xy = x$

$\Rightarrow \frac{dy}{dx} = x - 2xy$

$\Rightarrow \frac{dy}{dx} = x(1-2y)$

$\Rightarrow dy = x(1-2y)dx$

$\Rightarrow \frac{1}{1-2y}dy = x\,dx$

$\Rightarrow \int \frac{1}{1-2y}dy = \int x\,dx$

$\Rightarrow -\frac{1}{2}\ln(1-2y) = \frac{1}{2}x^2 + C$

$\Rightarrow e^{\ln(1-2y)} = e^{-x^2+c}$

$\Rightarrow 1-2y = ce^{-x^2}$

$\Rightarrow y = -\frac{ce^{-x^2}-1}{2}$

2. homogenous method

$f(kx, ky) = f(x,y)$

e.g. $\frac{dy}{dx} = \frac{x^2+y^2}{xy}$

check $\frac{k^2x^2+k^2y^2}{kxky} = \frac{x^2+y^2}{xy}$

let $v = \frac{y}{x}$ $\quad \frac{1}{v} = \frac{x}{y}$

$y = vx$

$\frac{dy}{dx} = x\frac{dv}{dx} + v$

• $\frac{dy}{dx} = \frac{\frac{x^2}{x^2}+\frac{y^2}{x^2}}{\frac{xy}{x^2}} = \frac{1+(\frac{y}{x})^2}{\frac{y}{x}}$

$= \frac{1+v^2}{v}$

$\Rightarrow x\frac{dv}{dx} + v = \frac{1+v^2}{v}$

$\Rightarrow x\frac{dv}{dx} + v = \frac{1}{v} + v$

$\Rightarrow x\frac{dv}{dx} = \frac{1}{v}$

$\Rightarrow v\,dv = \frac{1}{x}dx$

$\Rightarrow \int v\,dv = \int \frac{1}{x}dx$

$\Rightarrow \frac{1}{2}v^2 = \ln(x) + C$

$\Rightarrow v^2 = 2\ln|x| + C$

$v = \pm\sqrt{2\ln|x| + c}$

$\frac{y}{x} = \pm\sqrt{2\ln|x| + c}$

$y = \pm x\sqrt{2\ln|x| + c}$

3. Integrating factor

$\frac{dy}{dx} + P(x)y = Q(x)$

e.g. $\frac{dy}{dx} + 1y = x$ $\quad P(x) = 1$

$\Leftrightarrow \frac{dy}{dx} + P(x)y = 1$ ①

$\mu(x) = e^{\int P(x)dx}$

$\mu(x) = e^{\int 1\,dx}$

$= e^x$ ②

$\Rightarrow$ ① & ②

$\mu(x)\left[\frac{dy}{dx} + P(x)y = Q(x)\right]$

$\Rightarrow e^x\left[\frac{dy}{dx} + y = x\right]$

$\Rightarrow e^x\frac{dy}{dx} + ye^x = xe^x$

$\Rightarrow \frac{d}{dx}(e^x y) = xe^x$

$\Rightarrow \int \frac{d}{dx} e^x y = \int xe^x dx$ IBP

$\Rightarrow e^x y = xe^x - e^x + C$

$\Rightarrow y = x - 1 + \frac{c}{e^x}$

$= x - 1 + ce^{-x}$

• $\frac{dv}{dt} = g - \frac{c}{m}v$

$\frac{dv}{dt} + P(t)V = Q(t)$

$\Rightarrow \frac{dv}{dt} + (\frac{c}{m})V = g$

$\Rightarrow \mu(t) = e^{\int \frac{c}{m}dt}$

$= e^{\frac{c}{m}t}$

$\Rightarrow e^{\frac{c}{m}t}\left(\frac{dv}{dt} + (\frac{c}{m})v = g\right)$

$\Rightarrow e^{\frac{c}{m}t}\frac{dv}{dt} + \frac{c}{m}ve^{\frac{c}{m}t} = ge^{\frac{c}{m}t}$

$\Rightarrow \frac{d}{dt}e^{\frac{c}{m}t}v = ge^{\frac{c}{m}t}$

$\Rightarrow \int \frac{d}{dt}e^{\frac{c}{m}t}v = \int ge^{\frac{c}{m}t}dt$

$\Rightarrow e^{\frac{c}{m}t}v = \frac{m}{c}ge^{\frac{c}{m}t} + c'$

$\Rightarrow v = e^{-\frac{c}{m}t}\left[\frac{m}{c}ge^{\frac{c}{m}t} + c'\right]$

assume $V(0) = 0$

$v(0) = +\frac{m}{c}g + c' = 0$

$c' = -\frac{m}{c}g$

$\therefore v(t) = e^{-\frac{c}{m}t}\left[\frac{m}{c}ge^{\frac{c}{m}t} - \frac{m}{c}g\right]$

$= e^{-\frac{c}{m}t}\frac{m}{c}g\left[e^{\frac{c}{m}t} - 1\right]$

$= \frac{mg}{c}\left[1 - e^{-\frac{c}{m}t}\right]$

• linear ODE

- Laplace
- analytically solved
- $a_0(x)y + a_1(x)y' + \dots + a_n(x)y^{(n)} = b(x)$
- $Ly = f$

• Runge-Kutta Methods

△ Taylor series

$\frac{dy}{dx} = f(x,y)$

$\frac{d^{(n)}y}{dx^{n}} = f^{(n-1)}(x,y)$

$y_{i+1} = y_i + f(x_i,y_i)h$ Euler's Method

△ $y_{i+1} = y_i + f(x_i,y_i)h + \frac{f'(x_i,y_i)}{2!}h^2 + \dots + \frac{f^{(n-1)}(x_i,y_i)}{n!}h^n + O(h^{n+1})$

$\cong y_i + \left[f(x_i,y_i) + \frac{f'(x_i,y_i)}{2!}h + \dots + \frac{f^{(n-1)}(x_i,y_i)}{n!}h^{n-1}\right]h$

$= y_i + \phi(x_i,y_i,h)h$

△ $\phi(x_i,y_i,h) = f(x_i,y_i) + \frac{f'(x_i,y_i)}{2!}h + \dots + \frac{f^{(n-1)}(x_i,y_i)}{n!}h^{n-1}$

△ Runge-Kutta

acquire $\phi(x_i,y_i,h)$

w.o/ calculating higher-order term

$y_{i+1} = y_i + \phi(x_i,y_i,h)h$

$\Rightarrow \phi(x_i,y_i,h)$

$= a_1k_1 + a_2k_2 + \dots + a_{n-1}k_{n-1} + a_nk_n$

$k_1 = f(x_i,y_i)$

$k_2 = f(x_i+p_1h, y_i+q_{11}k_1h)$

$k_3 = f(x_i+p_2h, y_i+q_{21}k_1h+q_{22}k_2h)$

$\vdots$

$k_n = f(x_i+p_{n-1}h, y_i+q_{n-1,1}k_1h + q_{n-1,2}k_2h + q_{n-1,3}k_3h + \dots + q_{n-1,n-1}k_{n-1}h)$

a, p, q ⇒ constants

n = order ⇒ RKn method!

e.g.

△ $y = -0.5x^4 + 4x^3 - 10x^2 + 8.5x + 1$

$\frac{dy}{dx} = -2x^3 + 12x^2 - 20x + 8.5$

$f(0) = 1$

$f(0.5) = ?$

△ RK4:

$y_{i+1} = y_i + \phi(x_i,y_i,h)h$

$\phi(x_i,y_i,h)$

$= \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h$

$k_1 = f(x_i,y_i)$

$k_2 = f(x_i+\frac{1}{2}h, y_i+\frac{1}{2}k_1h)$

$k_3 = f(x_i+\frac{1}{2}h, y_i+\frac{1}{2}k_2h)$

$k_4 = f(x_i+h, y_i+k_3h)$

△ ∴

$k_1 = f(0,1) = 8.5$

$k_2 = f(0+0.25, 1+0.5\cdot8.5\cdot0.5) = 4.21875$

$k_3 = f(0+0.25, 1+0.5\cdot4.21875\cdot0.5) = 4.21875$

$k_4 = f(0.5, 1+4.21875\cdot0.5) = 1.25$

$y(0.5) = 1 + \frac{0.5}{6}(8.5 + 2\cdot4.21875 + 2\cdot4.21875 + 1.25)$

$= 3.21875 + O(h^5)$

• system of equations

$$\begin{cases} \frac{dy_1}{dx} = f_1(x, y_1, y_2, \dots, y_n) \\ \frac{dy_2}{dx} = f_2(x, y_1, y_2, \dots, y_n) \\ \vdots \\ \frac{dy_n}{dx} = f_n(x, y_1, y_2, \dots, y_n) \end{cases}$$

e.g.

△ $$\begin{cases} \frac{dy_1}{dx} = -0.5y_1 \\ \frac{dy_2}{dx} = 4 - 0.3y_2 - 0.1y_1 \end{cases}$$

$y_1 = 4$, $y_2 = 6$ @ $x = 0$

△ $\dot{y} = f(x)$

$$= \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix} = \begin{bmatrix} -0.5 & 0 \\ -0.1 & -0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

△ RK4

$y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h$

$k_1 = f(x_i, y_i)$

$k_2 = f(x_i+\frac{1}{2}h, y_i+\frac{1}{2}k_1h)$

$k_3 = f(x_i+\frac{1}{2}h, y_i+\frac{1}{2}k_2h)$

$k_4 = f(x_i+h, y_i+k_3h)$

△ $k_{1,1} = f_1(0,4) = -0.5\cdot4 = -2$

$k_{1,2} = f_2(0,6) = 4 - 0.3\cdot6 - 0.1\cdot4 = 1.8$

△ $k_{2,1} = f_1(0+\frac{1}{2}\cdot0.5, 4+\frac{1}{2}\cdot(-2)\cdot0.5)$

$= -0.5\cdot3.5 = -1.75$

$k_{2,2} = f_2(0+\frac{1}{2}0.5, 6+\frac{1}{2}\cdot(1.8)\cdot0.5)$

$= 4 - 0.3\cdot6.45 - 0.1\cdot3.5$

$= 1.715$

△ $k_{3,1} = f_1(0+\frac{1}{2}0.5, 4+\frac{1}{2}(-1.75)\cdot0.5)$

$= -0.5\cdot3.5625$

$= -1.78125$

$k_{3,2} = f_1(0+\frac{1}{2}\cdot0.25, 6+\frac{1}{2}(1.715)\cdot0.5)$

$= 4 - 0.3\cdot6.42875 - 0.1\cdot3.5625$

$= 1.715125$

△ $k_{4,1} = f_1(0+0.5, 4+(-1.78125)\cdot0.5)$

$= -0.5\cdot3.109375$

$= -1.5546875$

$k_{4,2} = f_2(0+0.5, 6+(1.715125\cdot0.5))$

$= 4 - 0.3\cdot6.8575625 - 0.1\cdot3.109375$

$= 1.63179375$

△ $y_{i+1} = y_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)h$

$y = \begin{bmatrix} 4 \\ 6 \end{bmatrix} + \frac{0.5}{6}\left(\begin{bmatrix} -2 \\ 1.8 \end{bmatrix} + 2\begin{bmatrix} -1.75 \\ 1.715 \end{bmatrix} + 2\begin{bmatrix} -1.78125 \\ 1.715125 \end{bmatrix} + \begin{bmatrix} -1.5546875 \\ 1.63179375 \end{bmatrix}\right)$

$= \begin{bmatrix} 3.115234 \\ 6.875670 \end{bmatrix}$ #

• solving high-order ODE

$$\begin{cases} \frac{d^2y}{dx^2} + 2\frac{dy}{dx} = x + y \\ y = 1 \quad \frac{dy}{dx} = 2 \quad \text{when } x = 0 \end{cases}$$

$$\begin{cases} \text{let } u = \frac{dy}{dx} \\ \frac{du}{dx} = \frac{d^2y}{dx^2} \end{cases}$$

$$\begin{cases} \Rightarrow \frac{du}{dx} + 2u = x + y \\ \Rightarrow \frac{du}{dx} = x + y - 2u \end{cases}$$

$$\begin{cases} \frac{dy}{dx} = u \\ \frac{du}{dx} = x + y - 2u \end{cases} \quad y(0) = 1, \; u(0) = 2$$

use this to propagate

$$\begin{cases} \frac{d^3y}{dx^3} + 2\frac{d^2y}{dx^2} + 3\frac{dy}{dx} = x + y \\ y = 1 \quad \frac{dy}{dx} = 2 \quad \frac{d^2y}{dx^2} = 3 \quad x = 0 \end{cases}$$

$$\begin{cases} \frac{dy}{dx} = u \\ \frac{d^2y}{dx^2} = v \end{cases} \quad \frac{d^2y}{dx^2} = \frac{du}{dx} = v \quad \frac{d^3y}{dx^3} = \frac{dv}{dx}$$

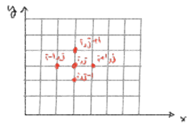
$$\begin{cases} \frac{dv}{dx} + 2v + 3u = x + y \\ \frac{dy}{dx} = u \\ \frac{du}{dx} = v \end{cases} \quad y(0) = 1, \; u(0) = 2, \; v(0) = 3, \; x = 0$$

use this to propagate

• PDE

• Laplace equation

$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0$



$\frac{\partial^2 T}{\partial x^2} = \frac{\left(\frac{T_{i+1,j} - T_{i,j}}{\Delta x}\right) - \left(\frac{T_{i,j} - T_{i-1,j}}{\Delta x}\right)}{\Delta x}$

$= \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{\Delta x^2}$

$\frac{\partial^2 T}{\partial y^2} = \frac{\left(\frac{T_{i,j+1} - T_{i,j}}{\Delta y}\right) - \left(\frac{T_{i,j} - T_{i,j-1}}{\Delta y}\right)}{\Delta y}$

$= \frac{T_{i,j+1} - 2T_{i,j} + T_{i,j-1}}{\Delta y^2}$

$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0$

$\Rightarrow \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{\Delta x^2} + \frac{T_{i,j+1} - 2T_{i,j} + T_{i,j-1}}{\Delta y^2} = 0$

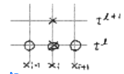
as $\Delta x = \Delta y$

$\Rightarrow T_{i+1,j} + T_{i-1,j} + T_{i,j+1} + T_{i,j-1} - 4T_{i,j} = 0$

$\Rightarrow T_{i,j} = \frac{T_{i+1,j} + T_{i-1,j} + T_{i,j+1}, T_{i,j-1}}{4}$

• Parabolic Equation

$k\frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}$



$k\frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}$

$\frac{\partial^2 T}{\partial x^2} = \frac{\frac{T^l_{i+1} - T^l_i}{\Delta x} - \frac{T^l_i - T^l_{i-1}}{\Delta x}}{\Delta x}$

$\frac{\partial T}{\partial t} = \frac{T^{l+1}_i - T^l_i}{\Delta t}$

$\Rightarrow k\frac{T^l_{i+1} - 2T^l_i + T^l_{i-1}}{\Delta x^2} = \frac{T^{l+1}_i - T^l_i}{\Delta t}$

$\Rightarrow T^{l+1}_i - T^l_i = \frac{k\Delta t}{\Delta x^2}(T^l_{i+1} - 2T^l_i + T^l_{i-1})$

$\Rightarrow T^{l+1}_i = T^l_i + \lambda(T^l_{i+1} - 2T^l_i + T^l_{i-1})$

$\lambda = \frac{k\Delta t}{\Delta x^2}$

e.g.

$\Delta x = 2$ cm

$\Delta t = 0.1$ s

$k = 0.835$ cm^2/s

$\lambda = 0.020875$

position

$$\begin{cases} T(0) = 100 \\ T(10) = 50 \end{cases} \quad @ \; \forall t$$

$T(x) = 0$ @ $t = 0$, $0 < x < 10$

$T^{l+1}_i = T^l_i + \lambda(T^l_{i+1} - 2T^l_i + T^l_{i-1})$

△ compute T @ $t = 0.1$ s, $x = 2$ cm

time

$T^1_1 = T^0_1 + 0.020875(T^0_2 - 2T^0_1 + T^0_0)$

position

$= 0 + 0.020875(0 - 2\cdot0 + 100)$

$= 2.0875$

△ compute T @ $t = 0.1$ s, $x = 4$ cm

$T^1_2 = T^0_2 + 0.020875(T^0_3 - 2T^0_2 + T^0_1)$

$= 0 + 0.020875(0 - 2\cdot0 + 0)$

$= 0$

△ compute T @ $t = 0.1$ s, $x = 6$ cm

$T^1_3 = T^0_3 + 0.020875(T^0_4 - 2T^0_3 + T^0_2)$

$= 0 + 0.020875(0 - 2\cdot0 + 0)$

$= 0$

△ compute T @ $t = 0.1$ s, $x = 8$ cm

$T^1_4 = T^0_4 + 0.020875(T^0_5 - 2T^0_4 + T^0_3)$

$= 0 + 0.020875(50 - 0 + 0)$

$= 1.0438$

△ Compute T @ $t = 0.2$ s, $x = 2$ cm

$T^2_1 = T^1_1 + \lambda(T^1_2 - 2T^1_1 + T^1_0)$

$= 2.0875 + 0.020875(0 - 2\cdot2.0875 + 100)$

$= 4.0878$

• convergence

$\lambda = \frac{k\Delta t}{\Delta x^2} \leq \frac{1}{2} \Rightarrow \Delta t \leq \frac{1}{2}\frac{\Delta x^2}{k}$

• $\frac{\partial T}{\partial t} = k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right)$

$$\begin{cases} \frac{\partial^2 T}{\partial x^2} = \frac{T^l_{i+1,j} - 2T^l_{i,j} + T^l_{i-1,j}}{\Delta x^2} \\ \frac{\partial^2 T}{\partial y^2} = \frac{T^l_{i,j+1} - 2T^l_{i,j} + T^l_{i,j-1}}{\Delta y^2} \\ \frac{\partial T}{\partial t} = \frac{T^{l+1}_{i,j} - T^l_{i,j}}{\Delta t} \end{cases}$$

$k\left(\frac{T^l_{i+1,j} - 2T^l_{i,j} + T^l_{i-1,j}}{\Delta x^2} + \frac{T^l_{i,j+1} - 2T^l_{i,j} + T^l_{i,j-1}}{\Delta y^2}\right)$

$= \frac{T^{l+1}_{i,j} - T^l_{i,j}}{\Delta t}$

$\Rightarrow T^{l+1}_{i,j} = T^l_{i,j} + \frac{k\Delta t}{\Delta x^2}(T^l_{i+1,j} - 2T^l_{i,j} + T^l_{i-1,j})$

$+ \frac{k\Delta t}{\Delta y^2}(T^l_{i,j+1} - 2T^l_{i,j} + T^l_{i,j-1})$

let $\Delta x = \Delta y$ $\quad \lambda = \frac{k\Delta t}{\Delta x^2} = \frac{k\Delta t}{\Delta y^2}$

$\Rightarrow T^{l+1}_{i,j} = T^l_{i,j} + \lambda(T^l_{i+1,j} - 2T^l_{i,j} + T^l_{i-1,j}$

$+ T^l_{i,j+1} - 2T^l_{i,j} + T^l_{i,j-1})$

• convergence

• 1-dimension :

$\lambda = \frac{k\Delta t}{\Delta x^2} \leq \frac{1}{2}$

• 2-dimension :

$\Delta t \leq \frac{1}{8}\frac{\Delta x^2 + \Delta y^2}{k}$

$\Downarrow$

if $\Delta x^2 = \Delta y^2$

$\Delta t \leq \frac{1}{4}\frac{\Delta x^2}{k}$

• Linear Programming

• Example problem

• A_1: origin 1 — costs c_{ij} — terminal 1 b_1, terminal 2 b_2, terminal 3 b_3

A_2: origin 2

$$\underset{x}{\text{minimize}}\ [c_{11}\ c_{12}\ c_{13}]\begin{bmatrix}x_{11}\\x_{12}\\x_{13}\end{bmatrix}+[c_{21}\ c_{22}\ c_{23}]\begin{bmatrix}x_{21}\\x_{22}\\x_{23}\end{bmatrix}$$

s.t. $\begin{cases}x_{11}+x_{12}+x_{13}=a_1\\x_{21}+x_{22}+x_{23}=a_2\end{cases}$ $\begin{cases}x_{11}+x_{21}=b_1\\x_{12}+x_{22}=b_2\\x_{13}+x_{23}=b_3\end{cases}$

$x_{ij}\succeq 0$

⇒ $\underset{x}{\text{minimize}}\ [c_{11}\ c_{12}\ c_{13}\ c_{21}\ c_{22}\ c_{23}]\begin{bmatrix}x_{11}\\x_{12}\\x_{13}\\x_{21}\\x_{22}\\x_{23}\end{bmatrix}$

s.t. $\begin{bmatrix}1&1&1&0&0&0\\0&0&0&1&1&1\\1&0&0&1&0&0\\0&1&0&0&1&0\\0&0&1&0&0&1\end{bmatrix}\begin{bmatrix}x_{11}\\x_{12}\\x_{13}\\x_{21}\\x_{22}\\x_{23}\end{bmatrix}=\begin{bmatrix}a_1\\a_2\\b_1\\b_2\\b_3\end{bmatrix}$, $x\succeq 0$

⇒ $\underset{x}{\text{minimize}}\ c^Tx$ s.t. $Ax=b$, $x\succeq 0$

• Transform to standard form

$\underset{x}{\text{minimize}}\ c^Tx$ s.t. $Ax\preceq b$, $x\succeq0$ ⇕ $\underset{x,y}{\text{minimize}}\ c^Tx$ s.t. $Ax+y=b$, $x\succeq0$, $y\succeq0$

$\underset{x}{\text{minimize}}\ c^Tx$ s.t. $Ax\succeq b$, $x\succeq0$ ⇕ $\underset{x,y}{\text{minimize}}\ c^Tx$ s.t. $Ax-y=b$, $x\succeq0$, $y\succeq0$

$\underset{x,y}{\text{minimize}}\ c^Tx+y$ s.t. $Ax+y=b$, $x\succeq0$ (no $y\succeq0$)

⇕ let $y=u-v$

$\underset{x,u,v}{\text{minimize}}\ c^Tx+u-v$ s.t. $Ax+u-v=b$, $x\succeq0$ $u\succeq0$ $v\succeq0$

$\underset{x,y,z}{\text{minimize}}\ c^Tx+y+z$ s.t. $Ax+y+z=b$, $y+z=d$, $x\succeq0$ $y\succeq0$ (no $z\succeq0$)

⇕ let $z=b-Ax-y$

$\underset{x,y}{\text{minimize}}\ c^Tx+y+b-Ax-y$ s.t. $y+b-Ax-y=d$, $x\succeq0$ $y\succeq0$

⇕ $\underset{x,y}{\text{minimize}}\ c^Tx+b-Ax$ s.t. $b-Ax=d$, $x\succeq0$ $y\succeq0$

• Solution — Simplex method

高中线性规划

• Solution — Basic Feasible Solution (BFS)

$\underset{x}{\text{minimize}}\ c^Tx$ s.t. $Ax=b$, $x\succeq0$

let $\tilde{x}$ be

1. $\tilde{x}\in D$, i.e. $A\tilde{x}=b$, $\tilde{x}\succeq0$
2. cols of A corresponding to nonzeros of $\tilde{x}$ are linearly independent

Theorem: $\tilde{x}$ is an extreme (rank(A)=m) iff $\tilde{x}$ is BFS

I. Basic/non-Basic variables

• for any BFS $\tilde{x}$, we rewrite:

$\tilde{x}=\begin{pmatrix}x_B\\x_N\end{pmatrix}$ B: nonzero

$A=(B, N)$ N: zero

$c=\begin{pmatrix}c_B\\c_N\end{pmatrix}$

• $Ax=b$ ⇕ $Bx_B+Nx_N=b$ ⇕ $Bx_B=b$ — rank(B) $=m_b$, $x_B=B^{-1}b$, $x_N=0^{n-m_b}$

• $Bx_B+Nx_N=b$

$\begin{cases}x_B=B^{-1}(b-Nx_N)\succeq0^{m_b}\\x_N\succeq0^{n-m_b}\end{cases}$

2. $\underset{x}{\text{minimize}}\ c^Tx$ s.t. $Ax=b$, $x\succeq0$ ⇔ $\underset{x}{\text{minimize}}\ c_B^Tx_B+c_N^Tx_N$ s.t. $Bx_B+Nx_N=b$, $x_B\succeq0^{m_b}$, $x_N\succeq0^{n-m_b}$

⇔ $\underset{x}{\text{minimize}}\ c_B^T(B^{-1}(b-Nx_N))+c_N^Tx_N$ s.t. $B^{-1}(b-Nx_N)\succeq0^{m_b}$, $x_N\succeq0^{n-m_b}$

• Simplex method: the Tableau

minimize x_1-x_2

s.t. $x_1-x_2\le2$, $x_1+x_2\le6$, $x_1\ge0$, $x_2\ge0$

$x_1=0$, $x_2=6$, $x_3=8$, $x_4=0$

⇕

minimize $p=x_1-x_2$

S.T. $x_1-x_2+s_1=2$, $x_1+x_2+s_2=6$, $x_1\ge0$ $x_2\ge0$, $s_1\ge0$ $s_2\ge0$

$-x_1+x_2=p$

$-p-x_1+x_2=0$

⇒ $p+x_1-x_2=0$

$$\begin{array}{cccccc} x_1 & x_2 & s_1 & s_2 & p & \\ \end{array}$$

$\begin{bmatrix}1&-1&1&0&0&2\\1&1&0&1&0&6\\1&-1&0&0&1&0\end{bmatrix}$ −2, 6

⇕

$\begin{bmatrix}2&0&1&1&0&8\\1&1&0&1&0&6\\1&-1&0&0&1&0\end{bmatrix}$

⇕

$\begin{bmatrix}2&\boxed{0}&\boxed{1}&1&0&\boxed{8}\\1&\boxed{1}&\boxed{0}&1&0&\boxed{6}\\2&0&0&1&1&6\end{bmatrix}$

maximize $p=5x+4y$

s.t. $3x+5y\le78$, $4x+y\le36$, $x,y\ge0$

⇕

maximize$_{x,y,s_1,s_2}$ $p=5x+4y$

$3x+5y+s_1=78$, $4x+y+s_2=36$, $x,y,s_1,s_2\ge0$

$$\begin{array}{cccccc} x & y & s_1 & s_2 & p & \\ \end{array}$$

$\begin{bmatrix}3&5&1&0&0&78\\4&1&0&1&0&36\\-5&-4&0&0&1&0\end{bmatrix}$ 26, 9

⇕

$\begin{bmatrix}3&5&1&0&0&78\\1&1/4&0&1/4&0&9\\-5&-4&0&0&1&0\end{bmatrix}$

⇕

$\begin{bmatrix}0&17/4&1&-3/4&0&51\\1&1/4&0&1/4&0&9\\0&-11/4&0&5/4&1&45\end{bmatrix}$ 12, 36

⇕

$\begin{bmatrix}0&1&4/17&-3/17&0&12\\1&1/4&0&1/4&0&9\\0&-11/4&0&5/4&1&45\end{bmatrix}$

⇕

$\begin{bmatrix}0&1&4/17&-3/17&0&12\\1&0&-\frac{1}{17}&\frac{5}{17}&0&6\\0&0&\frac{11}{17}&\frac{13}{17}&1&78\end{bmatrix}$

$x=6$, $y=12$, $P=78$

• Integer/Mixed-Integer Linear Programming

ILP/MILP

$\underset{x}{\text{minimize}}\ c^Tx$

s.t. $Ax=b$, $x\succeq0$, $x\in\{x\mid x_i\in\mathbb{Z}, i=1\dots l\}$

• Solution - Branch & Bound

• Branch: partition D into subsets (check objective value in subsets)

• Bound: acquire a bound w.o./ integer constraints

e.g.

$\underset{x}{\text{maximize}}\ 9x_1+5x_2+6x_3$

s.t. $6x_1+3x_2+5x_3\le10$, $-x_1+x_3\le0$, $x_i\in\{0,1\}$ $i=1,2,3$

⇕

$\underset{x}{\text{maximize}}\ 9x_1+5x_2+6x_3$

s.t. $6x_1+3x_2+5x_3\le10$, $-x_1+x_3\le0$, $x_i\le1$ $i=1,2,3$, $x_i\ge0$ $i=1,2,3$, $x_i\in\mathbb{Z}$ $i=1,2,3$

△ Bound:

• w.o./ $\mathbb{Z}$ constraints:

$x_1=1$, $x_2=1$, $x_3=0.2$

$P=15.2$ (UB)

• Rounded down solution

$x_1=1$, $x_2=1$, $x_3=0$

$p=14$ (LB)

△ Branch

• UB 15.2 (1, 1, 0.2)

LB 14 (1, 1, 0)

$x_3=0$ → feasible; $x_3=1$ → infeasible

UB 14 (1, 1, 0)

LB 14 (1, 1, 0)

no need to check hv

UB=LB

• Solutions to Optimization Problems.

• recall Duality

$\underset{x}{\text{minimize}}\ f_0(x)$

s.t. $f_i(x)\le0$, $h_i(x)=0$

△ Lagrangian:

$\mathcal{L}(x,\lambda,\nu)$

$=f_0(x)+\sum_{i=1}^m\lambda_if_i(x)+\sum_{i=1}^p\nu_ih_i(x)$

△ Lagrangian Dual function:

let

$g(\lambda,\nu)=\inf_{x\in D}\{\mathcal{L}(x,\lambda,\nu)\}$

$\underset{\lambda,\nu}{\text{maximize}}\ g(\lambda,\nu)$

s.t. $\lambda\succeq0$

△ let λ^*,ν^*, $d^*=g(\lambda^*,\nu^*)$ be optimal

$p^*\ge d^*$

△ Strong duality

Slater's condition

$\exists x\in\text{relint}\,D$

s.t. $f_i(x)<h_i$, $h_i(x)=0$

△ complementary slackness

$f_0(x^*)=g(\lambda^*,\nu^*)$

$=\inf_x f_0(x)+\sum_{i=1}^m\lambda_i^*f_i(x)+\sum_{i=1}^p\nu_i^*h_i(x)$

$\le f_0(x^*)+\sum_{i=1}^m\lambda_i^*f_i(x^*)+\sum_{i=1}^p\nu_i^*h_i(x^*)$

$\le f_0(x^*)$

$\lambda_i^*f_i(x^*)=0$

⇒ $\begin{cases}f_i(x^*)<0\to\lambda_i^*=0\\\lambda_i^*>0\to f_i(x^*)=0\end{cases}$

△ KKT optimality conditions

$\begin{cases}f_i(x^*)\le0\\h_i(x^*)=0\\\lambda^*\succeq0\\\lambda_i^*f_i(x^*)=0\\\nabla f_0(x^*)+\sum_{i=1}^m\lambda_i^*\nabla f_i(x^*)+\sum_{i=1}^p\nu_i^*\nabla h_i(x^*)=0\end{cases}$

this will be used later

• Descent Method

1. given $x^{(0)}\in\text{dom}\,f$
2. repeat
   1) descent direction Δx
   2) stepsize $t\ge0$
   3) $x^{k+1}=x^k+t^k\Delta x^k$
3. Stop $\|\Delta f(x)\|_2\le\eta$

• Line Search

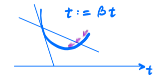
• Exact Line Search

$t=\underset{s\ge0}{\text{argmin}}\ f(x+s\Delta x)$

• Backtracking Line Search

1. given Δx^k, $f(x^k)$, $\alpha\in(0, 0.5)$, $\beta\in(0,1)$
2. set $t:=1$
3. while $f(x+t\Delta x)>f(x)+\alpha t\nabla f(x)^T\Delta x$

$t:=\beta t$



• Primal-dual Interior-point method

minimize $tf_0(x)+\phi(x)$

s.t. $Ax=b$

KKT:

$\begin{cases}Ax^*(t)=b\\t\nabla f_0(x^*(t))+\nabla\phi(x^*(t))+A^T\hat{\nu}=0\end{cases}$

⇒ $\begin{cases}Ax^*(t)=b\\t\nabla f_0(x^*(t))+\sum_{i=1}^m\frac{1}{-f_i(x^*(t))}\nabla f_i(x^*(t))+A^T\hat{\nu}=0\end{cases}$

⇒ $\begin{cases}Ax=b\\\nabla f_0(x)+\sum_{i=1}^m-\frac{1}{f_i(x)t}\nabla f_i(x)+A^T\nu=0\end{cases}$ $\frac{\hat{\nu}}{t}=\nu$

⇒ $\begin{cases}\nabla f_0(x)+\sum_{i=1}^m\lambda_i\nabla f_i(x)+A^T\nu=0\\-\lambda_if_i(x)=1/t\\Ax=b\end{cases}$

⇒ $r_t(x,\lambda,\nu)=\begin{bmatrix}\nabla f_0(x)+Df(x)^T\lambda+A^T\nu\\-\text{diag}(\lambda)f(x)-1/t\,\mathbf{1}\\Ax-b\end{bmatrix}$ — r_{dual}, r_{cent}, r_{primal}

⇒ $r_t(x+\Delta x,\lambda+\Delta\lambda,\nu+\Delta\nu)=\begin{bmatrix}\nabla^2f_0(x)+\sum_{i=1}^m\lambda_i\nabla^2f_i(x)&Df(x)^T&A^T\\-\text{diag}(\lambda)Df(x)&-\text{diag}(f(x))&0\\A&0&0\end{bmatrix}\begin{bmatrix}\Delta x\\\Delta\lambda\\\Delta\nu\end{bmatrix}=\begin{bmatrix}-r_{dual}\\-r_{cent}\\-r_{pri}\end{bmatrix}$

• Unconstrained problem

• Gradient descent

1. given $x_0\in\text{dom}\,f$
2. repeat
   - $\Delta x^k=-\nabla f(x)$
   - t: from line search
   - $x^{k+1}=x^k+t\Delta x^k$
3. until $\|\nabla f(x)\|\le\eta$

• Newton's method

1. given: $x^0\in\text{dom}\,f$
2. repeat:

$\Delta x^k=-\nabla^2f(x^k)^{-1}\nabla f(x^k)$

$\lambda^{k^2}=\nabla f(x^k)\nabla^2f(x^k)\nabla f(x^k)$

t: from line search

$x^{k+1}=x^k+t\Delta x$

3. until $\lambda^{k^2}/2\le\epsilon$

on Newton's Method

$f(x)$

⇒ $\hat{f}(x+v)=f(x)+\nabla f(x)^Tv+\frac{1}{2}v^T\nabla^2f(x)v$

⇒ $\nabla\hat{f}(x+v)=0$

⇒ $\nabla f(x)+\nabla^2f(x)v=0$

⇒ $v=-\nabla^2f(x)^{-1}\nabla f(x)$

• Equality-constrained Problem

minimize $f(x)$

s.t. $Ax=b$

recall

Slater's condition

$\exists x\in\text{relint}\,D$

s.t. $Ax=b$

KKT optimality condition

$f_i(x^*)\le0$

$h_i(x^*)=0$

$\lambda^*\succeq0$

$\lambda_i^*f_i(x^*)=0$

$\nabla f_0(x^*)+\sum_{i=1}^m\lambda_i^*\nabla f_i(x^*)+\sum_{i=1}^p\nu_i^*\nabla h_i(x^*)=0$

△ Here @ optimal $\tilde{x},\tilde{\nu}$

$\begin{cases}A\tilde{x}=b\\\nabla f(\tilde{x})+A^T\tilde{\nu}=0\end{cases}$

• Newton's Method

1. @ x

addition to x: $x+v$

∴ $x+v\in\text{dom}\,f$

$A(x+v)=b$

∴ $Av=0$ —①

2. $\hat{f}(x+v)$

$=f(x)+\nabla f(x)v+\frac{1}{2}v^T\nabla^2f(x)v$

∴ minimize $f(x)+\nabla f(x)v+\frac{1}{2}v^T\nabla^2f(x)v$

s.t. $Av=0$ —②

③ $\tilde{v}$ is optimal of ②

∴ $\exists w$ s.t. (Lagrange multiplier of ②)

$\nabla f(x)+\nabla^2f(x)\tilde{v}+A^Tw=0$

and recall $A\tilde{v}=0$

from KKT $\mathcal{L}(v,w)=\hat{f}(v)+w(Av)$

∴ $A\tilde{v}=0$

$\nabla f(x)+\nabla^2f(x)\tilde{v}+A^Tw=0$

⇒ $\begin{bmatrix}\nabla^2f(x)&A^T\\A&0\end{bmatrix}\begin{bmatrix}\tilde{v}\\w\end{bmatrix}=\begin{bmatrix}-\nabla f(x)\\0\end{bmatrix}$

∴ $\tilde{v}=\Delta x_{nt}$.

• Inequality-Constrained Problem

• Barrier Method

minimize $f_0(x)$

s.t. $f_i(x)\le0$, $Ax=b$

⇕

minimize $tf_0(x)+\phi(x)$

s.t. $Ax=b$

$\phi(x)=-\sum_{i=1}^m\log(-f_i(x))$

1. given $x^0\in\text{dom}\,f$

$t^0:=t^0>0$

$\mu>1$

$\epsilon>0$

2. repeat

a. solve

$\underset{x}{\text{minimize}}\ tf_0(x)+\phi(x)$

s.t. $Ax=b$

(get x^*, w/ t^k)

b. $x:=x^*(t)$

c. until $m/t<\epsilon$

d. $t:=\mu t$