Csongor Horváth: ○♠♣○ Exam Cheat Sheet ○♡◇○ Optimisation *MMXXII* 𝔚𝔚 ○♣♠○
Hosted in privat github repo *./horvathcso/Optimisation-exam-notes*. For acces contact: *sifuto2013@gmail.com*

feasible sol, constrain, LP, NLP, boundary, interior, global/local opt, standard form $S = \{x : Ax = b, x \geq 0\}$
$f : \mathbb{R}^n \to \mathbb{R}$ **Gradient** ∇f, **Hessian** $\nabla^2 f$ sim (Young), **Jacobian** ∇f^T, $f : \mathbb{R}^n \to \mathbb{R}^m$
Def. (Descent direction) *p dir of f in x, if* $\exists \epsilon > 0 :\ f(x + \alpha p) < f(x), \forall 0 < \alpha \leq \epsilon$
Rm. (Gradient) *p less than $90°$ from $-\nabla f(x)$ is descent, i.e* $\mathbf{p^T \nabla f(x)} < \mathbf{0}$

```
if x(k) loc opt: STOP
else: find p search dir, a step
do: x(k+1)=x(k)+a*p,           k++
```

Optimality conditions

Th. (1-ord nec) $\nabla f(x^*) = 0$ ***stationary***
Th. (2-ord nec) x^* *loc min* $\Rightarrow \nabla^2 f(x^*) \geq 0$
Th. (suf) $\nabla f(x^*) = 0$, $\nabla^2 f(x^*) > 0 \Rightarrow x^*$ *strict loc min*

Taylor series

$f(x + p) \approx f(x) + p\nabla f(x) + \frac{1}{2}p^T \nabla^2 f(x) p +$

Convexity

Def. (conv) $\forall x, y \in S :\ \alpha x + (1 - \alpha)y \in S$
$+ :\ f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$
Th. *Suppose convexity: loc min⇒global.*
Rm. *Constrained problem: If step size α feasible⇒ $\forall \beta < \alpha$ as well.*
Th. $\nabla^2 f \geq 0 \Rightarrow f$ *conv*, $\nabla^2 f > 0 \Rightarrow f$ *strict conv, f conv if above the tagent* $f(y) \geq f(x) + \nabla f(x)^T (y - x)$

Newton's Method for Nonlin eq

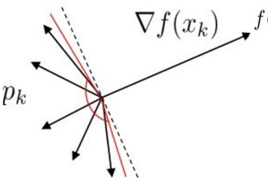
find f(x)=0: given x_0
do: $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$
$/x_{k+1} = x_k - f(x_k)\frac{x_k - k_{k-1}}{f(x_k) - f(x_{k-1})}$ deriv free

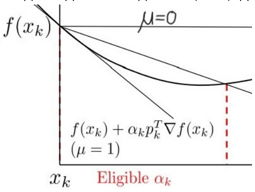
Th. *quadratic conv if $f' \neq 0$, f'' cont, x_0 close, ϵ_0 small. from Taylor:*
Biz. $\underbrace{x_* - x_{k+1}}_{\varepsilon_{k+1}} = \frac{-f''(\xi_k)}{2f'(x_k)}(\underbrace{x_* - x_k}_{\varepsilon_k})^2$ $\square$

find: g(x)=0: $(x_0,\ g : \mathbb{R}^n \to \mathbb{R}^n)$
do: $x_{k+1} = x_k - \nabla g(x_k)^{-T} g(x_k)$
$p \Leftarrow g(x_k + p) \approx g(x_k) + \nabla g(x_k)^T p = 0$

Newton's Method for optim

Idea: $\nabla f(x)$ neccesery cond.
$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$ quadratic
Sup.lin $x_{k+1} = x_k + p_k$, if $lim \frac{||p_k - p_{N_k}||}{||p_k||} = 0$
glob. strat. if fails: no conv, sing Hessian
Desc dir: $p_k^T \nabla f_{x_k} = -\nabla f^T [\nabla^2 f]^{-1} \nabla f < 0$
$a)$:angle condition $-\frac{p_k^T \nabla f(x_k)}{||p_k|| ||\nabla f(x_k)||} \geq \epsilon_1 > 0$
$b)$:"grad related" $M \geq ||p_k|| \geq \epsilon_2 ||\nabla f(x_k)||$





c) Armijo "sufficient decrease" $\mu \in (0, 1)$: $f(x_k + \alpha_k p_k) \leq f(x_k) + \mu \alpha_k p_k^T \nabla f(x_k)$,
d) Backtracking line search-"not too small"

while: $||\nabla f(x_k)|| \geq \epsilon$
do: $\nabla^2 f(x_k) + D_1 = LDL^T$,
p from: $(LDL^T)p = -\nabla f(x_k)$
do: line search $x_{k+1} = x_k + \alpha_k p_k$

Th. $f : \mathbb{R}^n \to \mathbb{R}, \{x : f(x) \geq f(x_0)\}$ *bound*, ∇f *Lipsch cont*, $a) - d) \Rightarrow lim \nabla f(x_k) = 0$
Biz. 1)f bound below, 2)$lim \exists$ 3) $lim \alpha_k ||\nabla f(x_k)||^2 = 0$ 4)$1 \geq \alpha_k \geq \epsilon ||\nabla f(x_k)||^2$ $\square$

Steepest-Descent method

$p_k = -\nabla f(x_k)$, α_k from line search.
$a), b)$ satifsied trivially + conv rate linear.
With exact line search $x_k \perp x_{k+1}$
Lem. $e_{k+1} \leq \left(\frac{cond(Q) - 1}{cond(Q) + 1}\right)^2 e_k$, *if* $f = x^T Q x - c^T x$, *Q poz def*, $e_k = f x_k - f x_*$

Quasi-Newton Methods

given: x_0, B_0 **if**: x_k opt STOP
lin srch $x_{k+1} = x_k + \alpha_k(-B_k^{-1} \nabla f(x_k))$
s_k, y_k **and** $B_{k+1} = B_k + \ldots$ compute

$\nabla^2 f$ and $(\nabla^2 f)^{-1}$ expensive to compute
idea: $B_k p = -\nabla f(x_k) \to p = B_k^{-1} \nabla f(x_k)$
• be symm (like Hess) • esay update
• $B = I$ steepest desc, • $B = Hess$ Newton
Secant:$1D$ opt:$x_{k+1} = x_k - \frac{f'(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$
Def. (Secant cond) $B_k \cdot (s_k) = y_k$
$s_k = x_k - x_{k-1}$, $y_k = \nabla f(x_k) - \nabla f(x_{k-1})$
$B_{k+1} = B_k + smth$ •secant, •$\mathcal{O}(n^2)$ update + inverz, •symm, •pos def
sym rnk-1 $smth = \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}$
BFGS rnk-2 $smth = \frac{y_k y_k^T}{y_k^T s_k} - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k}$
Broyden cl: $\Phi = 0$ BFGS, $\Phi = 1$ DFP
$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + \Phi \cdot (s_k^T B_k s_k) v_k^T$, where $v_k = \frac{y_k}{y_k^T s_k} - \frac{B_k s_k}{s_k^T B_k s_k}$
Lem. (BFGS) *B_k poz def, B_{k+1} poz def* $\Leftrightarrow y_k^T s_k > 0$
Th. (Broyden cl) • $S = \{f(x) \leq f(x_0)\}$ *bnded* • *$f, \nabla f, \nabla^2 f$ cont in S* •*$\nabla^2 f$ poz def* •$p_k = B_k^{-1} - \nabla f(x_k)$, $B_0 = I$, $0 \leq \Phi < 1$ •α_k : *Armijo&*$\exists 0 < \nu < 1 :\ p_k^T \nabla f(x_k + \alpha_k p_k) \geq \nu p_k^T \nabla f(x_k) \Rightarrow$ $\mathbf{lim x_k = x_*}$, *x_* unique glob min, conv rate superlinear.*

Constrained

min $f(x)$, *s.t* : $g_i(x) = / \geq 0, i \in E$
loc opt: x_*: $f(x_*) \leq f(x), \forall x$ feasible $\&|x - x_*| < \epsilon \Rightarrow$ interior same cond
Bound point opt: no feasible&descent dir.
$p \nabla f(x_*) \geq 0$ for all feas. p
Asm: loc opt, 2 differ, constr quali, full rank

Lin eq

min$f(x)$ $Ax = / \geq b$ convex space

=: p ∈nullspace → Z nullspace matrix: $N(A) = Im(Z)$
Reformul:x feasible:$min_v \varphi(v) = f(x + Zv)$
$\nabla \varphi = Z^T \nabla f(x), \nabla^2 \varphi = Z^T \nabla^2 f(x) Z$
So we can use the opt conditions convex space
With Lagrange multipliers: $\nabla f(x_*) = A^T \lambda_* \Leftrightarrow$ 1-ord condition.
repr: change in opt val to unit change in b

Lr func: $\mathcal{L}(x, \lambda) = f(x) - \lambda^T (Ax - b)$
stat point $(x, \lambda) :\ \nabla \mathcal{L}(x, \lambda) = 0$

≥: polytope/polyhedron
only active constr relevant to opt. den:$\hat{A}, \hat{b}$
necces: *min*$f(x)$; $\hat{A}x = \hat{b} \Rightarrow$
1-ord: $Z^T \nabla f(x_*) = 0$ or $\exists \hat{\lambda_*} \geq 0 :$ $\nabla f(x_*) = \hat{A}^T \hat{\lambda_*}$ 2-ord:$Z^T \nabla^2 f Z$ pos semi.
Complementary slackness: $\lambda_*^T (Ax_* - b) = 0, \lambda \geq 0$, So $\lambda_i = 0 \Leftrightarrow i$ non-active rows in A
necc: 2-ord:$Z^T \nabla^2 f(x_*) Z$ pos semi, 1-ord: $\exists \lambda_* \geq 0 : \nabla f(x_*) = A^T \lambda_*, \lambda_*^T (Ax_* - b) = 0$
suff: *1-ord* + strict compl slack: $a_i^T x_* - b_i = 0 / \lambda_{*i} = 0$ and *2-ord:* pos def

Non-Lin eq

min$f(x)$, $g_i(x) = / \geq 0$, feas curve
$Z(x_k)$ nullspace of Jacobian $\nabla g(x_k)^T$
1-ord: stat point lag: $\mathcal{L}(x, \lambda) = f(x) - \sum \lambda_i g_i(x)$ cond: $\nabla f = \sum \lambda_i \nabla g$
necc for =: 1-ord $\nabla_x \mathcal{L}(x_*, \lambda_*) = 0$,
2-ord: $Z(x_*) \nabla_{xx}^2 \mathcal{L}(x_*, \lambda_*) Z(x_*)$ pos semi.
suff: $g_i(x_*) = 0$ + pos definite the end.
KKT: feasible(= / ≥), no-feasible decrease dir, compl slack(≥), pos lagrange mult(≥)

Methods

lin eq: reform + Newton: $-\nabla^2 \varphi^{-1} \cdot \nabla \varphi$
$x_{k+1} = x_k + Z[Z^T \nabla^2 f(x_k) Z]^{-1} Z^T \nabla f(x_k)$
active-set: "guess" active constraints set. W active constraints set. ITERATE: *opt check:* •no active •optimal to active: λ •if $\lambda \geq 0$ stop else left negative
search dir descent p respect to W
step len: α: $f(x_k + \alpha p) < f(x_k)$, $\alpha \leq \hat{\alpha}$
update, **SQP***
solving constr opt as sequence of unconstr

barrier: *min*f, $g \geq 0$ $\mu_k \to 0$
Sequence, nice small is hard, and holds conv
Log bar func. $f(x) - \mu \sum_{i=1}^m \log(g_i(x))$
Inverse barrier func $f(x) + \mu \sum_{i=1}^m \frac{1}{g_i(x)}$
penalty meth: *min*f, $g = 0$, $\mu_k \to \infty$
min $f + \mu \frac{1}{2} \sum g_i^2$ *problem may:* no f on feas region.

LP

Lin func: $f = c^Tx$, $\nabla f = c$, **Quadr func**: $f = \frac{1}{2}x^TQx - b^Tx$, $\nabla f = Qx - b$, $\nabla^2 f = Q$

Condition numb: $cond(A) = ||A||\cdot||A^{-1}||$

Def. (Standard) $min\ cx\ s.t.\ Ax = b,\ x \geq 0$

All LP can be transformed to standard form

•feasible region: $\mathcal{P}$ polytope*/ polyhedron

• $\mathcal{P}^* = conv$(vertex) *Representation thm*

Def. (basic solution) $Ax = b$,& *clmns* A_j *for* $x_j \neq 0$ *linearly independent, basis* B

Def. (BFS) + $x \geq 0$ - $x_j \neq 0$ *basic var*

• x extreme point ⇔ BFS • adj - ∃ edge

•adj basis-1col dif •adj basis ∼ adj ext. pnt

St. *if* c *bounded in* $\mathcal{P}$, *than* $\exists x^*$ *BFS opt.*

Th. x *extrm* ⇔ *BFS*, $x \in S$

Biz. ⇒ $x = (B, N)$, B may not square, suppose B lin. dep. $Bp = 0$, $x_B \geq 0$, $\exists \epsilon > 0 : x \pm \epsilon p$

⇐ supp. $x = (x_B, x_N)$ not extreme, $\exists (y_B, y_N), (z_B, z_N)$: $\alpha y + (1-\alpha)z = x$, $y_N = z_N = 0$, & $\exists B^{-1}$ so $x = y = z$↯ □

Def. (degenerate v) ∃ *basic part of sol* = 0

Def. (redundant constrain) *slack variable*

Def. (Direction of unbondness) $x + \alpha d \in S$, $\forall \alpha \geq 0 \Rightarrow Ad = 0$ *for standard form*

Th. (Repr) $\forall x \in S:\ x = d + \sum \alpha_i v_i$, $V = \{v_i\}$ *vertices*

Th. *If* c *on* S *bounded* ⇒ ∃ *opt. BFS*

Simplex

idea: ∃ opt ext point=BFS ⇔ basis of A

Iterate in adj vertices (bases) always decrease until opt.

Df: *(non-)basis & entering/leaving vari*

Df: $x_B = B^{-1}b$ cur sol, $z = c_B^TB^{-1}b$ cur obj

---

given: basis B, BFS $x_B = B^{-1}b \geq 0$

if opt?: ?$c_N = c_N - N^TB^{-T}c_B \geq 0$□

else: x_t: $\hat{c}_t < 0$ entering (Bland)

$s = argmin_i\{\frac{\hat{b}_i}{\hat{a}_{i,t}} : \hat{a}_{i,t} > 0\}$ **if** $\nexists a_{i,t}$ unbnd

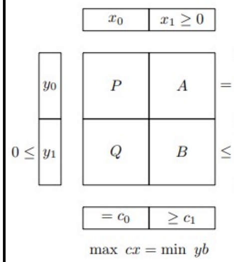
else update: $x_b -= \hat{A}_t x_t,\ z += \hat{c}_t x_t$

$x_B = B^{-1}b - \alpha(B^{-1}N)_t,\ x_N = 0 + \alpha\delta_t$

init sol: find smhow/ 2-phase/ punish slack

Unbounded: recognize in step 2 if $\forall a_{i,t} < 0$

Degeneracy: ∃0 valued basic variable, may be $\alpha = 0$, baj cycling: Bland rule / perturbation = lexicographic mthd



$$\begin{array}{ll} min & c^Tx \\ s.t. & Ax \geq b \\ & x \geq 0 \\ \hline max & b^Ty \\ s.t. & A^Ty \leq c \\ & y \geq 0 \end{array}$$

Th. (weak duality) *If* x *primal feasible,* y *dual feasible:* $cx \geq by$

Th. (strong duality) *If prim or dual* ∃ *bounded opt:* ∃*opt other&* $mincx = maxyb$

Th. (Complementary slackness) x^*, y^* *prim, dual opt sol:* $x^T(c - A^Ty) = 0$, *and if* x, y *feasible and = holds:* x, y *opt sol.*

So either $x_j = 0$ or $(A^Ty)_j = c_j\ \forall j$

Null&Range space

Def. $A \in \mathbb{R}^{m\times n}$, $m \leq n$: $\mathcal{N}(A) = \{p \in \mathbb{R} : Ap = 0\}$, $\mathcal{R}(A^T) = \{q \in \mathbb{R} :\ q = A^Tx\}$

$dim(\mathcal{N}(A)) = n - rk(A)$, $\mathcal{N}(A) \perp \mathcal{R}(A^T)$

Sequential quadratic programming: $minf(x)$, $g(x) = 0$, 1-ord lin apprx of:$\nabla\mathcal{L}(x, \lambda) = 0$ Newton:$\begin{pmatrix} x_{k+1} \\ \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} + \begin{pmatrix} p_k \\ v_k \end{pmatrix}$,

p_k, v_k from: $\nabla^2\mathcal{L}(x_k, \lambda_k)\begin{pmatrix} p_k \\ v_k \end{pmatrix} = -\nabla\mathcal{L}(x_k, \lambda_k) . = \begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x_k, \lambda_k) & -\nabla g(x_k) \\ -\nabla g(x_k)^T & 0 \end{pmatrix}\begin{pmatrix} p_k \\ v_k \end{pmatrix} = \begin{pmatrix} -\nabla_x\mathcal{L}(x_k, \lambda_k) \\ g(x_k) \end{pmatrix}$.

1-ord cond in quadratic from: $\begin{array}{ll} \text{minimize} & q(p) = \frac{1}{2}p^T\left[\nabla^2_{xx}\mathcal{L}(x_k, \lambda_k)\right]p + p^T\left[\nabla_x\mathcal{L}(x_k, \lambda_k)\right] \\ \text{subject to} & \left[\nabla g(x_k)\right]^Tp + g(x_k) = 0, \end{array}$