Quasi-Newton Methods

<u>BFGS</u> (Broyden, Fletcher, Goldfarb, Shanno)

Consider (again) $m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p$

for B_k SPD and B_k approx. Hessian

Proceeding as for Newton: $p_k = -B_k^{-1} \nabla f_k$

$$x_{k+1} = x_k + \alpha_k p_k$$

where α_k satisfies Wolfe cond.s

In the next step we want new model

$m_{k+1}(p) = f_{k+1} + \nabla f_{k+1}^T p + \frac{1}{2} p^T B_{k+1} p$

s.t. $\nabla m_{k+1}(0) = \nabla f_{k+1}$ (obvious)

$$\nabla m_{k+1}(-\alpha_k p_k) = \nabla f_k \quad \} \rightarrow$$

$$\nabla f_{k+1} - \alpha_k B_{k+1} p_k = \nabla f_k \iff$$

$$B_{k+1} \alpha_k p_k = \nabla f_{k+1} - \nabla f_k$$

Set $x_{k+1} - x_k =: s_k \quad y_k := \nabla f_{k+1} - \nabla f_k$:

$$B_{k+1} s_k = y_k \quad \text{(second eq. / cond.)}$$

B_{k+1} SPD $\rightarrow s_k^T B_{k+1} s_k = s_k^T y_k > 0$

poses constraints on LS. Wolfe or SW cond.s guarantee this curvature cond.

curv. cond: $\nabla f_{k+1}^T p_k \geq c_2 \nabla f_k^T p_k \Rightarrow (s_k = \alpha_k p_k)$

$$\nabla f_{k+1}^T s_k \geq c_2 \nabla f_k^T s_k$$

$(\nabla f_{k+1} - \nabla f_k)^T s_k \geq (c_2 - 1) \nabla f_k^T s_k > 0$

since $c_2 < 1$ and $\nabla f_k^T p_k < 0$ $(\alpha_k > 0)$

Many possible SPD choices B_{k+1}

For BFGS:

let $H_k = B_k^{-1}$ (also SPD) $\rightarrow H_{k+1} y_k = s_k$

and $\min_H \| H - H_k \| \rightarrow H_{k+1}$

subject to $H^T = H$ and $H y_k = s_k$

Take $\| \bar{G}_k^{1/2} (H - H_k) \bar{G}_k^{1/2} \|_F$ min where

$\bar{G}_k$ average Hessian $\bar{G}_k = \int_0^1 \nabla^2 f(x_k + \tau \alpha_k p_k) \, d\tau$

(BFGS) Then $H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$

where $\rho_k = (y_k^T s_k)^{-1}$

How to pick H_0?

exact Hessian or good approx. at x_0
(e.g. by finite diff.s)

constant times I (poss. 1)

(p. 143) Use $\tilde{H}_0 = I$ to get p_k, compute y_k, s_k

and then set H_0 to $\dfrac{y_0^T s_0}{y_0^T y_0} I$ before

computing the update to get H_1

The BFGS update of B_k (corr. to update H_k)

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

(However, requires Chol. decomp $\rightarrow O(n^3)$ cost)
poss. to update Chol. factors directly $\rightarrow O(n^2)$ cost.

If B_k SPD then B_{k+1} SPD (assuming α_k following WC or SWC)

$$z^T H_{k+1} z = \quad (z \neq 0)$$

$$\underbrace{z^T(I - \rho_k s_k y_k^T)}_{w^T} H_k \underbrace{(I - \rho_k y_k s_k^T) z}_{w} + \rho_k z^T s_k s_k^T z =$$

$$\underbrace{w^T H_k w}_{>0 \text{ if } w \neq 0} + (y_k^T s_k)^{-1} \underbrace{(z^T s_k)^2}_{>0 \text{ if } z^T s_k \neq 0}$$

if $z^T s_k = 0$ then $\overset{\to w = z}{w \neq 0}$ and $z^T H_{k+1} z > 0$

if $z^T s_k = 0$ then $z^T H_{k+1} z = z^T H_k z > 0$

An alternative for BFGS is SR1 method that satisfies some remarkable properties (later), and often approx. Hessian better than BFGS.

SR1 req. only rank 1 update. Since it may not maintain pos. def. better used with TR than with LS. (see Alg. 6.2)

$$B_{k+1} = B_k + \sigma v v^T \qquad \sigma = \pm 1$$

$$y_k = (B_k + \sigma v v^T) s_k = B_k s_k + \sigma v v^T s_k \iff$$

$\leftarrow$ scalar

$$y_k - B_k s_k = v \, (\sigma v^T s_k)$$

Hence $v = \delta (y_k - B_k s_k) \rightarrow$

$$(y_k - B_k s_k) = \sigma \delta^2 \left(s_k^T (y_k - B_k s_k)\right)(y_k - B_k s_k)$$

If $y_k - B_k s_k = 0 \rightarrow v = 0$: $B_{k+1} = B_k$

If not, in principle: $\begin{cases} \delta = |s_k^T (y_k - B_k s_k)|^{-\frac{1}{2}} \\ \sigma = \text{sign}\left(s_k^T (y_k - B_k s_k)\right) \end{cases}$

if δ small problem $\rightarrow$

if $|s_k^T (y_k - B_k s_k)| \geq r \|s_k\| \, \|y_k - B_k s_k\|$

for a small $r \in (0,1)$, e.g., $r = 10^{-8}$ (book)

then update is done, otherwise $B_{k+1} = B_k$

<u>Theo 6.1</u>: Suppose $f: \mathbb{R}^n \to \mathbb{R}$ defined as $f(x) = \frac{1}{2} x^T A x + b^T x$, where A SPD. Then for any starting point x_0 and any symmetric starting matrix H_0, the iterates $\{x_k\}$ from SR1 with

$$\begin{cases} H_{k+1} = H_k + \dfrac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k} \\ p_k = -H_k \nabla f_k, \quad x_{k+1} = x_k + p_k \end{cases}$$

converge to x^* in n steps, provided

$(s_k - H_k y_k)^T y_k \neq 0$ for all k.

Moreover, if n steps are performed and the search directions p_i linearly indep., then $H_n = A^{-1}$.

Proof: (see book, simple by induction showing $H_{k+1} y_j = s_j \quad j = 0..k$)

Theorem 6.2: Suppose f twice cont. diff. and Hessian bnd and Lip. cont. in nbhood of point x^*. Let $\{x_k\}$ be any seq. of iterates s.t. $x_k \to x^*$ for some $x^* \in \mathbb{R}^n$. In addition, suppose for all k
$|s_k^T(y_k - B_k s_k)| \geq r \|s_k\| \|y_k - B_k s_k\|$ for some $r \in (0,1)$
and that the s_k are uniformly indep. (form space of max dim). Then $\lim_{k\to\infty} \|B_k - \nabla^2 f(x^*)\| = 0$

If f twice cont. diff. and the level set $L = \{x \mid f(x) \leq f(x_0)\}$ is convex, and Hessian $G(x)$ satisfies $0 < m \leq z^T G(x) z \leq M < \infty$ for all $x \in L$ and all $z \in \mathbb{R}^n$. In addition if B_0 SPD. Then $\{x_k\}$ generated by BFGS implemented by Alg. 6.1 converges to the minimizer x^* of f.

Note that assumptions are significantly stronger than usual.

Theo.
If f twice cont. diff. and $\{x_k\}$ generated by BFGS converge to x^* s.t. $\sum_{k=1}^{\infty} \|x_k - x^*\| < \infty$ and Hessian $G(x)$ Lip. cont at x^*, (locally)

$$\|G(x) - G(x^*)\| \le L \|x - x^*\| \text{ for all } x \text{ near } x^*$$

$(L > 0)$, then $x_k \to x^*$ superlinearly.

~~Hen~~ Proof, see book.

Large-scale Unconstrained Optimization

Inexact Newton Methods

Solve $\nabla^2 f_k \, p_k = -\nabla f_k$ or $B_k p_k = -\nabla f_k$

use CG for approx. sol.

residual : $r_k = \nabla^2 f_k \, p_k + \nabla f_k$

termination cond. $\|r_k\| \leq \eta_k \|\nabla f_k\|$

for seq. $\{\eta_k\}$ w. $0 < \eta_k < 1$ $(\forall k)$

↳ forcing seq.

Assume $\nabla^2 f(x)$ cont in nbhood men. x^* and pos. def. at x^*

Use a forcing seq. $\eta_k \leq \eta$ $\eta \in [0,1)$

For x suff. close to x^*, $\|\nabla^2 f(x)^{-1}\| \leq L$

$$p_k = (\nabla^2 f_k)^{-1} (r_k + \nabla f_k)$$

$$\|p_k\| \leq L \left(\|\nabla f_k\| + \underbrace{\|r_k\|}_{\leq \eta \|\nabla f_k\|}\right) \leq 2L \|\nabla f_k\|$$

$$\cancel{\nabla f_{k+1} = \nabla f(x_k + \alpha p_k) = \nabla f_k + \nabla^2 f_k \, p_k + \int}$$

mistake book

$$\nabla f_{k+1} = \nabla f(x_k + \alpha_k p_k) =$$

$$\nabla f_k + \underbrace{\nabla^2 f_k \, p_k}_{r_k - \nabla f_k} + \underbrace{\int_0^1 [\nabla^2 f(x_k + t p_k) - \nabla^2 f(x_k)] \, p_k \, dt}_{\| \cdots \| = o(\|p_k\|)}$$

$$= r_k + o(\|\nabla f_k\|)$$

$$\|\nabla f_{k+1}\| \le \eta_k \|\nabla f_k\| + o(\|\nabla f_k\|)$$
$$\le (\eta_k + o(1))\, \|\nabla f_k\|$$ (too many braces in bk)

$o(1)$ term $\to 0$, ~~so~~ assume $o(\|p_k\|)$ term smaller than $(1-\eta)/2$

$$\|\nabla f_{k+1}\| \le \frac{1+\eta}{2} \|\nabla f_k\|$$ (close enough to x^*)

~~&~~ Start close enough to x^* $\to$ have this each step.

$$\nabla f_k = \nabla^2 f_k(x^*)(x_k - x^*) + o(\|x_k - x^*\|)$$
$$\|\nabla f_k\| \le \|\nabla^2 f(x^*)\| \, \|e_k\| + o(\|e_k\|)$$

Above $\dfrac{\|f_{k+1}\|}{\|f_k\|} \le \eta_k + o(1)$ or

$$\left(\frac{\|\nabla f_{k+1}\|}{\|\nabla f_k\|} \le \frac{1+\eta}{2} \right)$$

So, for $\eta_k \to 0$ we get

$$\lim_{k\to\infty} \frac{\|f_{k+1}\|}{\|\nabla f_k\|} = 0$$ superlinear conv.

SPD $\nabla^2 f$ implies $\nabla f = 0 \iff x_k = x^*$

of grad. to zero (implies superlin conv x_k to x^*)

If $\nabla^2 f$ Lip. cont. near x^*, then

$$\|\nabla f_{k+1}\| = O(\|\nabla f_k\|^2)$$ (quadr. conv.)

if $\eta_k = O(\|\nabla f_k\|)$

$\nabla^2 f$ lip cont in nbhood x^* and pos. def x^*

Forcing seq. $\{\eta_k\}$ with $\eta_k < \eta \in [0,1)$

For x suff. close to x^* $\|\nabla^2 f(x)^{-1}\| \leq L$

$p_k = (\nabla^2 f_k)^{-1}(r_k + \nabla f_k)$ $\quad \|r_k\| \leq \eta_k \|\nabla f_k\|$

$\|p_k\| \leq L(\|\nabla f_k\| + \|r_k\|)$

$\leq 2L\|\nabla f_k\|$

$\nabla f_{k+1} = \nabla f(x_k + \alpha_k p_k) =$

$\nabla f_k + \nabla^2 f_k p_k + \underbrace{\int_0^1 [\nabla^2 f(x_k + t p_k) - \nabla^2 f(x_k)] p_k \, dt}$

$\| \quad \| \leq \frac{1}{2}\gamma \|p_k\|^2$

$\|\nabla^2 f(x_k + t p_k) - \nabla^2 f(x_k)\| \leq \gamma \|t p_k\| = \gamma t \|p_k\|$

These results $\Rightarrow$ Theo. 7.2

sugg. for η_k: $\begin{cases} \eta_k = \min(1/2, \|\nabla f_k\|^{1/2}) \\ \eta_k = \min(1/2, \|\nabla f_k\|) \end{cases}$

Line Search Newton CG for

~~[illegible]~~ $\cdot\ B_k p = -\nabla f_k$

Choose x_0 ; $k=0$; tol ; comp $\nabla f(x_0)$

while $\|\nabla f_k\|$ ~~$\leq$~~ $>$ tol,

$\varepsilon_k =$ ~~min~~ $\eta_k \|\nabla f_k\|$

$z_0 = 0$; $r_0 = +\nabla f_k$; $d_0 = -r_0$; $j=0$

while $\|r_j\| > \varepsilon_k$,

if $d_j^T B_k d_j \leq 0$ (neg. curv.)

if $j=0$ then $p_k = -\nabla f_k$

else $p_k = z_j$

$\alpha_j = r_j^T r_j / d_j^T B_k d_j$

$z_{j+1} = z_j + \alpha_j d_j$

$r_{j+1} = r_j + \alpha_j B_k d_j$

if $\|r_{j+1}\| > \varepsilon_k$

$\beta_{j+1} = r_{j+1}^T r_{j+1} / r_j^T r_j$

$d_{j+1} = -r_{j+1} + \beta_{j+1} d_j$

end while

$x_{k+1} = x_k + \alpha_k p_k$ (where α_k sat.

Wolfe/Armijo cond. and $\alpha_k = 1$

when possible $\rightarrow$ backtracking)

end

Essentially CG alg. with add test for B_k indef and special update w. direction of negative curvature.

Choice of η_k as suggested above.

For SPD $B_k = \nabla^2 f_k$ $\quad z_j \to p_k^N$

! B_k needed only for matrix-vector product. So, matrix-free impl. poss.

$$\nabla^2 f_k \cdot d \simeq \frac{\nabla^2 f(x_k + hd) - \underset{\text{known}}{\nabla f_k}}{h}$$

for small but not too small h;

essentially strike balance between accuracy in exact arithmetic $h \simeq 0$ and cancellation (numerical error) if $\dfrac{\| \nabla^2 f(x_k + hd) - \nabla^2 f_k \|}{\| \nabla f_k \|}$ too small.

If B_k too ill-cond. LS-CG not so effective.

Alg. 7.2

p. 171

Trust-Region CG

Use modified CG to solve/minimize approximately for TR step p

$m_k(p) = f_k + g_k^T p + \frac{1}{2} p^T B_k p$

(B may be indefinite) $\|p\| \leq \Delta$

* Use conv. criterion similar to CG-LS → forcing seq. η_k (ε_k)
* Stop on direction of negative curv. find min on bnd TR
* $z_0 = 0$

$$z_1 = -\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T B_k \nabla f_k} \nabla f_k$$

if $\|z_1\| \leq \Delta$ → Cauchy pt.

subseq. steps improve

if $\|z_1\| \geq \Delta$ → Cauchy pt and stop

* $\|z_j\| > \|z_{j-1}\| > \cdots$

So, can always stop at TR bnd.

Prop. follows from prop. of CG and choice $z_0 = 0$.

Theo 7.3 $\{z_j\}$ from Alg 7.2 satisfies

$0 = \|z_0\| < \|z_1\| < \cdots < \|p_k\| \leq \Delta_k$

The CG iter. can be preconditioned to accelerate convergence.
Often used are variants of Inexact Modified Cholesky.

The choice of step when neg curv. is discovered, can be poor.

As alternative we can use Lanczos method for tridiag. reduction/proj. of B_k (as in CG) and then use this directly for minimization.

$B_k p = -\nabla f_k$

Iterate as for CG $(z_0 = 0)$:

$$B_k Q_j = Q_j T_j + \beta_j q_{j+1} e_j^T$$

$* \rightarrow q_j = \pm r_j / \|r_j\|$

$r_0 = \nabla f_k$

$$T_j = Q_j^T B_k Q_j \quad \text{and} \quad Q_j^T Q_j = I \; *$$

Find minimum for TR problem from $R(Q_j)$: $\rightarrow p = Q_j w$

$$\min_{w \in \mathbb{R}^j} \; f_k + \nabla f_k^T Q_j w + \tfrac{1}{2} w^T Q_j^T B_k Q_j w$$

$$\text{subj } \|Q_j w\| = \Delta$$

Simplify: $q_1 = \nabla f / \|\nabla f_k\|$ and $Q_j^T Q_j = I$

$$\rightarrow \nabla f_k^T Q_j w = \|\nabla f_k\| e_1^T w$$

$$Q_j^T B_k Q_j = T_j \quad \text{tridiag.}$$

$$\|Q_j w\| = \Delta \implies \|w\| = \Delta$$

$$\min_{w \in \mathbb{R}^j} \; f_k + \|\nabla f_k\| e_1^T w + \tfrac{1}{2} w^T T_j w$$

Since T_j is tridiag and small

we can easily factorize $T_j + \lambda I$ and follow the (nearly) exact TR model/subproblem solution approach of section 4.3.

Limited-memory BFGS

Inverse updated BFGS:

$x_{k+1} = x_k - \alpha_k H_k \nabla f_k$

$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T$ where

$\rho_k = \frac{1}{y_k^T s_k}$, $V_k = I - \rho_k y_k s_k^T$

$s_k = x_{k+1} - x_k$, $y_k = \nabla f_{k+1} - \nabla f_k$

H_k in general dense → expensive for large problems.

Solution: store H_k implicitly (approx.) by keeping m pairs $\{s_i, y_i\}$

$H_k \nabla f_k$ can then be computed efficiently

Keep m pairs $\{s_i, y_i\}$ $i = k-m, \dots, k-1$

H_k^0 given (simple SPD matrix)

$H_k = H_k^{(k)}$

$H_k^{\,} = V_{k-1}^T H_k^{(k-1)} V_{k-1} + \rho_{k-1} s_{k-1} s_{k-1}^T =$

$= V_{k-1}^T \left(V_{k-2}^T H_k^{k-2} V_{k-2} + \rho_{k-2} s_{k-2} s_{k-2}^T \right) V_{k-1}$

$+ \rho_{k-1} s_{k-1} s_{k-1}^T$

$V_k = I - \rho_k y_k s_k^T$

$$= V_{k-1}^T V_{k-2}^T \cdots V_{k-m}^T H_k^{(0)} V_{k-m} \cdots V_{k-2} V_{k-1}$$
$$+ \rho_{k-m} (V_{k-1}^T \cdots V_{k-m+1}^T) s_{k-m} s_{k-m}^T V_{k-m+1} \cdots V_{k-1}$$
$$+ \vdots$$
$$+ \rho_{k-1} s_{k-1} s_{k-1}^T$$

↓ Can be implemented eff. using recursion.
→ Alg. 7.4 (explicitly unrolled)

Poss. choice $H_k^{(0)} = \gamma_k I$ where

$$\gamma_k = \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}$$

L-BFGS
Choose x_0, $m > 0$
$k = 0$
while $\|\nabla f_k\| >$ tol an $k <$ maxit

  Choose $H_k^{(0)}$ — $H_k^{(0)}$ may vary per step

  $p_k = -H_k \nabla f_k$ (using approx repr.)

  $x_{k+1} = x_k + \alpha_k p_k$ (α_k sat's Wolfe cond's)

  if $k > m$
    Discard $\{s_{k-m}, y_{k-m}\}$

  Compute and save $s_k = x_{k+1} - x_k$
    $y_k = \nabla f_{k+1} - \nabla f_k$

  $k = k+1$
end

Compact Representation of BFGS updating

For TR we need limited memory BFGS representation for B_k itself.

*1 $$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

$$B_{k+1} s_k = y_k \rightarrow B_k s_{k-1} = y_{k-1}$$

$$B_{k+1} = B_k\left(I - \frac{s_k s_k^T}{s_k^T B_k s_k}\right) B_k + \frac{y_k y_k^T}{y_k^T s_k}$$

$$= \cancel{\left(B_{k-1}\left(I - \frac{s_{k-1} s_{k-1}^T}{s_{k-1}^T B_{k-1} s_{k-1}}\right) B_{k-1}\right)}$$

$$= \left(B_{k-1}\left(I - \frac{s_{k-1} s_{k-1}^T}{s_{k-1}^T B_{k-1} s_{k-1}}\right) B_{k-1} + \frac{y_{k-1} y_{k-1}^T}{y_{k-1}^T s_{k-1}}\right)\left(I - \cdots\right)\left(\quad\right)$$

$$+ \frac{y_k y_k}{y_k^T s_k} \longrightarrow$$

*2 $$B_k = B_0 - \begin{bmatrix} B_0 S_k & Y_k \end{bmatrix} \begin{bmatrix} S_k^T B_0 S_k & L_k \\ L_k^T & -D_k \end{bmatrix}^{-1} \begin{bmatrix} S_k^T B_0 \\ Y_k^T \end{bmatrix}$$

$$S_k = [s_0 \cdots s_{k-1}] \qquad Y_k = [y_0 \cdots y_{k-1}]$$

$$(L_k)_{ij} = \begin{cases} s_{i-1}^T y_{j-1} & \text{if } i > j \\ 0 & \text{otherwise} \end{cases}$$

$$D_k = \mathrm{diag}\left(s_0^T y_0, \ldots, s_{k-1}^T y_{k-1}\right)$$

If B_0 SPD and $\{s_i, y_i\}_{i=0}^{k-1}$ sat. $s_i^T y_i > 0$ forall i then B_k (*1) $=$ B_k (*2) and $\begin{bmatrix} \times & \cdot \\ \cdot & \times \end{bmatrix}^{-1}$ exists.

(note inverse exists because B_0 SPD and assumed $s_0^T y_0 > 0$

$$B_1 = B_0 - [B_0 s_0 \;\; y_0] \begin{bmatrix} s_0^T B_0 s_0 & 0 \\ 0 & -s_0^T y_0 \end{bmatrix}^{-1} \begin{bmatrix} s_0^T B_0 \\ y_0^T \end{bmatrix}$$

$$= B_0 - [B_0 s_0 \;\; y_0] \begin{bmatrix} (s_0^T B_0 s_0)^{-1} s_0^T B_0 \\ -(s_0^T y_0)^{-1} y_0^T \end{bmatrix}$$

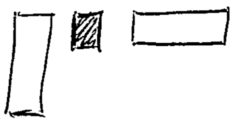
$$= B_0 - \left[B_0 s_0 s_0^T B_0 / (s_0^T B_0 s_0)^{-1} - y_0 y_0^T (s_0^T y_0) \right]$$

(verified for $k = 1 \rightarrow B_1$)

Assume true for $k = 1 \ldots m-1$

$$B_{m-1} = B_0 - [B_0 S_{m-1} \;\; Y_{m-1}] \begin{bmatrix} S_{m-1}^T B S_{m-1} & L_{m-1} \\ L_{m-1}^T & -D_{m-1} \end{bmatrix}^{-1} \begin{bmatrix} S_{m-1}^T B_0 \\ Y_{m-1}^T \end{bmatrix}$$

$B_m =$

$\delta_k I +$ 

For $k>m$ its, keeping only m pairs of vectors:

$$B_0 = \delta_k I = \gamma_k^{-1} I \quad \left(\gamma_k = \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}}\right)$$

$$S_k = [s_{k-m}\ s_{k-m+1} \cdots s_{k-1}] \quad Y_k = [y_{k-m} \cdots y_{k-1}]$$

$$B_k = \delta_k I - [\delta_k S_k\ Y_k]\begin{bmatrix}\delta_k S_k^T S_k & L_k \\ L_k^T & -D_k\end{bmatrix}^{-1}\begin{bmatrix}\delta_k S_k^T \\ Y_k^T\end{bmatrix}$$

$$(L_k)_{ij} = \begin{cases} s_{k-m-1+i}^T\, y_{k-m-1+j} & \text{if } i>j \\ 0 & \text{otherwise}\end{cases}$$

$$D_k = \mathrm{diag}(s_{k-m}^T y_{k-m} \cdots s_{k-1}^T y_{k-1})$$

$$\cancel{B_k = \delta I - UEU^T}$$

$$B_k = \delta I - W F^{-1} W^T$$

$$F = \begin{bmatrix}\delta S^T S & L \\ L^T & -D\end{bmatrix} =$$

$$\underbrace{\begin{bmatrix}\delta^{1/2}(S^TS)^{1/2} & 0 \\ \delta^{-1/2} L^T (S^TS)^{-1/2} & -G^{1/2}\end{bmatrix}}_{L}\underbrace{\begin{bmatrix}\delta^{1/2}(S^TS)^{1/2} & \delta^{-1/2}(S^TS)^{-1/2}L \\ 0 & G^{1/2}\end{bmatrix}}_{R}$$

$$G = +\frac{1}{\delta} L^T (S^TS)^{-1} L + D$$

$$B_k = \delta I - (WR^{-1})(L^{-1}W^T) = \delta I - UV^T$$

$$B_k^{-1} = \delta^{-1} I + UZV^T \Rightarrow$$

$$\cancel{I} - \delta^{-1}UV^T + \delta UZV^T - UZV^TUV^T = \cancel{I}$$

$$U(-\delta^{-1} I + \delta Z - ZV^TU)V^T = 0$$

$$-\delta^{-1} I + \delta Z - ZV^TU = 0 \Leftrightarrow -I + \delta^2 Z - \delta Z V^T U = 0$$

$$Z(\delta^2 I - \delta V^T U) = I \Leftrightarrow Z = (\delta^2 I - \delta V^T U)^{-1}$$

$\hookrightarrow 2k \times 2k$

So, using ~~implicit~~ this easy to compute inverse for B_k, we can easily implement Dogleg or other methods for TR (approx) solution.

Minimize $\frac{1}{2}\sum_i r_i^2(x) = f(x)$

$$\frac{\partial f}{\partial x_k} = \sum_i r_i(x) \cdot \frac{\partial r_i}{\partial x_k}$$

$$= r^T \left(\frac{\partial r}{\partial x}\right)_k \quad (k^{th} \text{ column})$$

$$= (J^T r)_k \quad k^{th} \text{ coeff.}$$

$$J_{ij} = \left(\frac{\partial r_i}{\partial x_k}\right)$$

$$\nabla f = J^T r$$

$$H_{jk} = H_{kj} = \frac{\partial f}{\partial x_j \partial x_k} = \frac{\partial}{\partial x_j}\left(\sum_i r_i \frac{\partial r_i}{\partial x_k}\right)$$

$$= \sum_i \underbrace{\frac{\partial r_i}{\partial x_j} \cdot \frac{\partial r_i}{\partial x_k}} + r_i \frac{\partial^2 r_i}{\partial x_j \partial x_k}$$

$$\sum_i \frac{\partial r_i}{\partial x_j} \cdot \frac{\partial r_i}{\partial x_k} = (J)_k^T (J)_j$$

$$\nabla^2 f = J^T J + \sum_i r_i (\nabla^2 r_i)$$

$\nabla^2 r_i$ often difficult and/or very expensive to compute. (sometimes exact impossible)

If $r \simeq 0$ close to solution, or at least $\|r\| \ll \|J^T J\|$ then we neglect $\sum_i r_i (\nabla^2 r_i) \rightarrow$

Gauss-Newton method/model

Local model	$f(x_k+p) \simeq m_k(p) = f_k + r^T J p + \frac{1}{2} p^T J^T J p$
	Newton step: $J^T J p = - J^T r$
	Normal eq.s for $Jp \simeq -r \rightarrow \min_p \|Jp + r\|_2$
In general J not square	$\rightarrow$ Linear Least Squares
	$J^{m \times n}$ $\quad m > n$
	$J : \mathbb{R}^n \rightarrow R(J) \subseteq \mathbb{R}^m$
	$r = r_1 + r_2$, $\quad r_1 \in R(J)$, $\quad r_2 \in R(J)^\perp$
Assume col.s J indep, otherwise $\tilde{n} < n$	QR-decomp J: $J = \underset{\text{orthog}}{Q^{m \times m}} \cdot \underset{\text{upper tri}}{R^{m \times n}}$
	$R = \begin{bmatrix} \underbrace{\diagdown}_{n \times n} \\ \hline 0 \end{bmatrix}$
	$J = m\begin{array}{c} \begin{array}{cc} n & m-n \end{array} \\ \left[\, Q_1 \mid Q_2 \,\right] \end{array} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1$
$Q_1 Q_1^T$ orthog.	where $Q_1^T Q_1 = I_n$ $\quad R(Q_1) = R(J)$
proj. in $\langle \cdot, \cdot \rangle_2$	$J p_1 = Q_1 Q_1^T r$ (part we can solve for)
assoc. with $\|\cdot\|_2$	$(I - Q_1 Q_1^T) r$ (residual)
	$Q_1^T (Jp - r) = 0 \rightarrow$ no comp. of residual in space $R(J)$ (otherwise better solution exists)

$Q_1^T(Q_1 R_1 p - r) = 0 \iff R_1 p = Q_1^T r$

Alternatively $\mathcal{J} = U \Sigma V^T_{n\times n}$

$$\underbrace{[U_1 | U_2]}_{m\times m} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{n} \\ & O & \end{bmatrix}}_{m\times n} \begin{bmatrix} V^T \end{bmatrix} =$$

$$\underset{m\times n}{U_1} \underset{n\times n}{\Sigma_n} \underset{n\times n}{V^T}$$ (~~allows for $\sigma_1 \geq \ldots \geq \sigma_p >$~~

allows for some (or ~~σ_p~~ all) of σ_i to be zero.

Let $p \leq \min(m, n)$ s.t.

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p > \sigma_{p+1} = \cdots = \sigma_n = 0$

$\mathcal{J} = [u_1 \cdots u_p] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_p^T \end{bmatrix}$

$U_p \Sigma_p V_p^T$

also works if $m < n$!

$\mathcal{R}(\mathcal{J}) = s\{u_1, \ldots, u_p\}$

$p = V_p \Sigma_p^{-1} U_p^T r$ pseudo inverse

(least norm solution / gen. inverse)

Such solutions or approx. can be computed in a number of ways.

Damped GN: ~~[illegible]~~

p is (approx) sol. of LS. problem

$x_{k+1} = x_k + \alpha_k p_k$

where α_k typically found by backtracking alg.

→ first $\alpha_k = 1$ if Armijo cond. sat. accept, otherwise reduce α_k, until Armijo cond sat.

→ Book: if Wolfe conds satisfied and

J full rank (throughout L) and

$\sigma_{min}(J) \geq \gamma > 0$
$\sigma_{max}(J) \leq \beta \; (< \infty)$

then convergence guaranteed (std line search arg. Zoutendyk's condition and bound on $\cos\theta_k$)

If $r(x^*) = 0 \rightarrow$ quadr. conv.
otherwise $\rightarrow$ (super) linear conv.

Even if J^TJ poss. singular, conv. can often be guaranteed if some, more elaborate, weaker cond. are satisfied.

$$p_{GN} = -V_p \Sigma_p^{-1} U_p^T r$$

$$-\nabla f^T p_{GN} = -r^T J p$$

$$= +r^T U_p \Sigma_p \cancel{V_p^T V_p} \Sigma_p U_p^T r$$

$$= +(U_p^T r)^T \Sigma_p^2 (U_p^T r)$$

unif. bnd from 0 if $p \geq 1$

$(U_p^T r) \geq \rho > 0$ and $\sigma_p \geq \bar{\sigma} > 0$
↳ $\sigma_{min} > 0$

$$\begin{cases} \min_p \frac{1}{2}\|J_k p + r_k\|_2^2 \\ \|p\|_2 \le \Delta_k \end{cases}$$

We can guarantee pos. def of approx. Hessian : $(J^T J + \lambda I)$ <u>for any $\lambda > 0$</u>

This leads essentially to trust region approach or slight variant:

<u>Levenberg-Marquardt method</u>

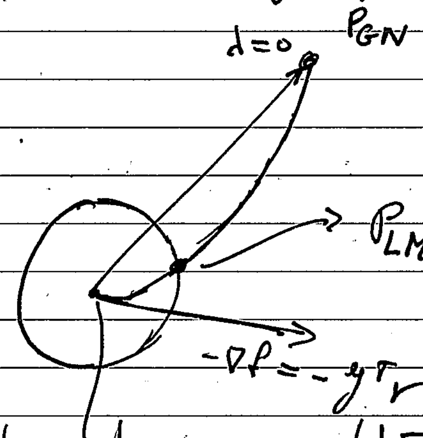
$$p_{LM} = -(J^T J + \lambda I)^{-1} J^T r$$

(Solve $(J^T J + \lambda I)p = -J^T r$)

Choice of λ important and similar to TR radius

~~Note~~ Assume $(J^T J)^{-1} J^T r = p_{GN}$ and $\|p_{GN}\| > \Delta$

Consider $p(\lambda) \leftarrow (J^T J + \lambda I)p(\lambda) = -J^T r$



$\lambda \to \infty$: $(\lambda I)p = -J^T r \Rightarrow p = -\frac{1}{\lambda} J^T r$

$\hookrightarrow$ LM gives solution balancing / weighting (neg) gradient and (Gauss) Newton step.

Compute in LS sense

$$p_{LM} = \arg\min_{p} \frac{1}{2} \left\| \begin{bmatrix} J \\ \sqrt{\lambda} I \end{bmatrix} p + \begin{bmatrix} r \\ 0 \end{bmatrix} \right\|_2^2$$

For implementation details, see pp. 259 and further

A convergence proof largely follows the std proof for TR methods.

~~Com~~ Comparison of Damped GN and LM.

$$\underset{n\times p}{}\ \underset{p\times p}{}\ \underset{p\times n}{}$$

$J = U_l \Sigma_l V_l^T \quad l \le \min(m,n)$

$\sigma_l > \sigma_{l+1} = 0$

GN $\rightarrow$ $p_{GN} = -V_l \Sigma_l^{-1} U_l^T r$ ← assuming we compute least norm solutions of J singular

$(J^T J = V_l \Sigma_l^T \Sigma_l V_l^T, \; J^T r = V_l \Sigma_l^T U_l^T r$

$J^T J^{+} J^T r = V_l \Sigma_l^{-2} \cancel{V_l^T V_l} \Sigma_l U_l^T r$

$\quad = V_l \Sigma_l^{-1} U_l^T r)$

Assume line search needed $\rightarrow$

$p_k = \alpha_k p_{GN} = -\alpha_k V_l \Sigma_l^{-1} U_l^T r$

If some σ_i (very) small, typically α_k very small

Note $p_k = V_l \begin{pmatrix} \alpha_k & & \\ & \ddots & \\ & & \alpha_k \end{pmatrix} \Sigma_l^{-1} U_l^T r$

$$= -\sum_{i=1}^{l} v_i \cdot \alpha_k \cdot \frac{u_i^T r}{\sigma_i}$$

α_k small, all comps damped equally strongly (indep of need) | LS damping applied component-wise

LM.

$\cancel{J^TJ = U_q \Sigma_q V_q^T}$ $\quad J = U_q \Sigma_q V_q^T$

$\cancel{J^TJ + \lambda I}$ $\quad q = \min(m, n)$

$\Sigma_q = \mathrm{diag}(\sigma_1, \dots, \sigma_q)$ and

possibly some $\sigma_i = 0$

(assume $\sigma_1 \geq \sigma_2 \geq \cdots$)

$$J^TJ = V_q \Sigma_q^T \Sigma_q V_q^T$$

$$\hookrightarrow \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_q^2 \end{pmatrix} \text{ if } m < n$$

$$J^TJ + \lambda I = V_q (\Sigma_q^T \Sigma_q + \lambda I) V_q^T$$

$$\begin{pmatrix} \sigma_1^2 + \lambda & & \\ & \ddots & \\ & & \sigma_q^2 + \lambda \end{pmatrix}$$

$$(J^TJ + \lambda I)^{-1} J^T r =$$

$$- V_q (\Sigma_q^T \Sigma_q + \lambda I)^{-1} \cancel{V_q^T V_q} \Sigma_q^T U_q^T r$$

$$= - \sum_{i=1}^{q} v_i \frac{\sigma_i}{\sigma_i^2 + \lambda} u_i^T r$$

$$= - \sum_{i=1}^{q} v_i \frac{\sigma_i^2}{\sigma_i^2 + \lambda} \cdot \frac{u_i^T r}{\sigma_i}$$

$\hookrightarrow$ damping coeff.s

damping of indiv. components depends on magn. of λ vs σ_i

$\sigma_i^2 \gg \lambda$ very little damping

$\sigma_i^2 = \lambda \rightarrow \frac{1}{2}$

$\sigma_i^2 \ll \lambda \rightarrow$ very strong damping

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subj. to} \quad \begin{cases} c_i(x) = 0 & i \in \mathcal{E} \\ c_i(x) \geq 0 & i \in \mathcal{I} \end{cases}$$

f, c_i all smooth, real valued on subset of $\mathbb{R}^n$

$\mathcal{I}, \mathcal{E}$ finite index sets

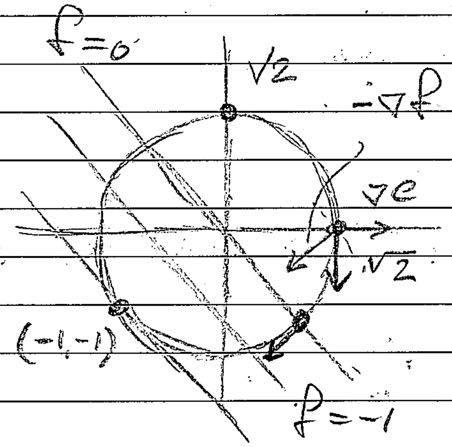
$$\Omega = \{x \in \mathbb{R}^n \mid c_i(x) \geq 0, i \in \mathcal{I}, \; c_i(x) = 0, i \in \mathcal{E}\}$$

↳ Ω bounded by piecewise smooth curves/surfaces

Active set $A(x)$ at any feasible x consists of set of set of indices of constr. satisfied as equality constraints.

$$A(x) = \mathcal{E} \cup \{i \in \mathcal{I} \mid c_i(x) = 0\}$$

Ex. 12.1 $\quad \min\; x_1 + x_2 \quad$ s.t. $x_1^2 + x_2^2 - 2 = 0$



$x_1 + x_2 = c \rightarrow x_1 = c - x_2$
$\qquad\qquad\qquad\quad x_2 = c - x_1$

x : stay on constr. curve

$c(x) = 0$ and

$$c(x+s) = c(x) + \nabla c(x)^T s + \underbrace{\mathcal{O}(\|s\|^2)}_{=0}$$

for small s : $\nabla c^T s = 0$

To reduce f : $\nabla f^T s < s$

Unless ∇f and ∇c parallel, we can find direction s which reduces f along constraint curve/surface.

Altern. if such s does not exist, we satisfy cond.s for (constr.) local extremum
→ Lagrange multiplier method.

$$L = f(x) + \lambda c(x) \rightarrow$$

stat. point L : $$\begin{cases} \frac{\partial L}{\partial x} = \nabla_x f + \lambda \nabla_x c = 0 \\ \frac{\partial L}{\partial \lambda} = c(x) = 0 \end{cases}$$

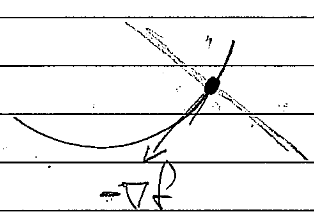
First cond : $\nabla f = -\lambda \nabla c$ <u>parallel</u>

$$\begin{cases} 1 + \lambda(2x_1) = 0 \\ 1 + \lambda(2x_2) = 0 \\ x_1^2 + x_2^2 = 2 \end{cases} \qquad \begin{array}{l} 2\lambda x_1 = -1 \quad \lambda x_1 = -\frac{1}{2} \\ \lambda x_2 = -\frac{1}{2} \\ x_1^2 + x_2^2 = 2 \end{array}$$

$$\underbrace{\lambda^2 x_1^2}_{1/4} + \underbrace{\lambda^2 x_2^2}_{1/4} = 2\lambda^2 \rightarrow \lambda^2 = \frac{1}{4} \rightarrow \lambda = \pm\frac{1}{2}$$

$\lambda = \frac{1}{2} \rightarrow x_1 = x_2 = -1$ → max
$\lambda = -\frac{1}{2} \rightarrow x_1 = x_2 = 1$ → min
} stat. points

$\min x_1 + x_2$ s.t. $x_1^2 + x_2^2 \leq 2 \rightarrow 2 - x_1^2 - x_2^2 \geq 0$



Still need $\nabla f^T s < 0$
but only $c_1 + \nabla c^T s \geq 0$ (to first order)

If x in interior of region $c(x) \geq 0$
any small enough step satisf. 2 crit.
So, only need $\nabla f^T s < 0$

If $\nabla f \neq 0 \rightarrow s = -\alpha \nabla f$

If x on bnd situation as before.

$h = f + \lambda c$

Opt. (first order → stat. pt):

$\nabla_x h = \nabla f + \lambda \nabla c = 0$ for some $\lambda \geq 0$

and $\lambda c(x) = 0$ (either $\lambda = 0$ or $c = 0$)

↳ complementarity condition

$\{$ $c = 0$ λ nonzero ~~can be anything~~, constr. active and need ∇c parallel to ∇f

$c > 0$ $\lambda = 0$ so, that constr. has no effect on (local) solution

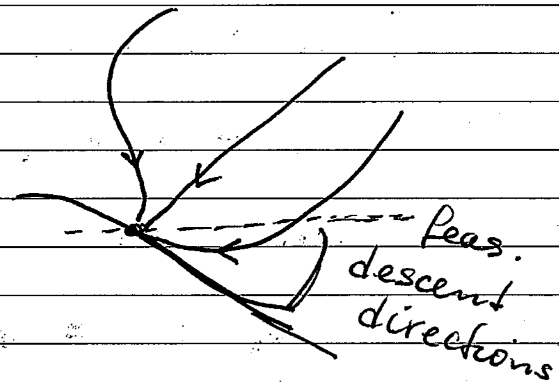
Need robust ~~good~~ methods to come up with useful search directions taking both obj. function and constraints into account.

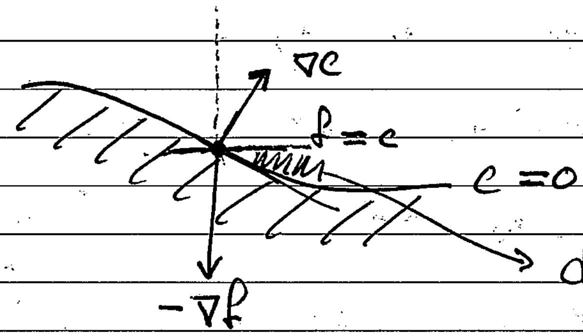
$\{z_k\}$ feasible sequence approaching x if $z_k \in \Omega$ for all k suff. large and $z_k \to x$.

We can define local solution as point x^* s.t. all feasible sequences approaching x^* have $f(z_k) \geq f(x^*)$ for all suff. large k.

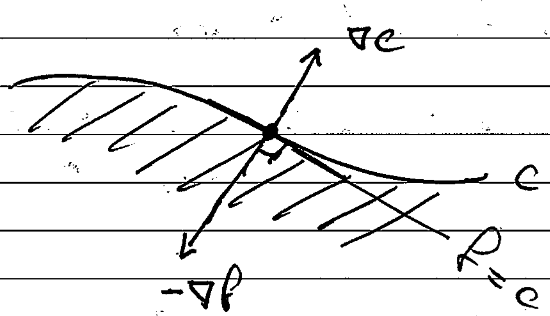
$\min f(x) \quad \text{subj } c_1 \geq 0$

$$\begin{cases} \nabla f - d\,\nabla c = 0 & d \geq 0 \\ c \geq 0 & \end{cases}$$





$\nabla f - d\,\nabla c = 0$ for some $d > 0$



Def 12.2

Ω feasible region

The vector d is called a tangent (vector) to Ω at x if there are a feasible seq. $\{z_k\}$ approaching x and a seq. of pos. scalars $\{t_k\}$ w. $t_k \to 0$ s.t.

$$\lim_{k\to\infty} \frac{z_k - x}{t_k} = d$$

A set F is a cone if $x \in F \Rightarrow \alpha x \in F$ $\forall \alpha > 0$

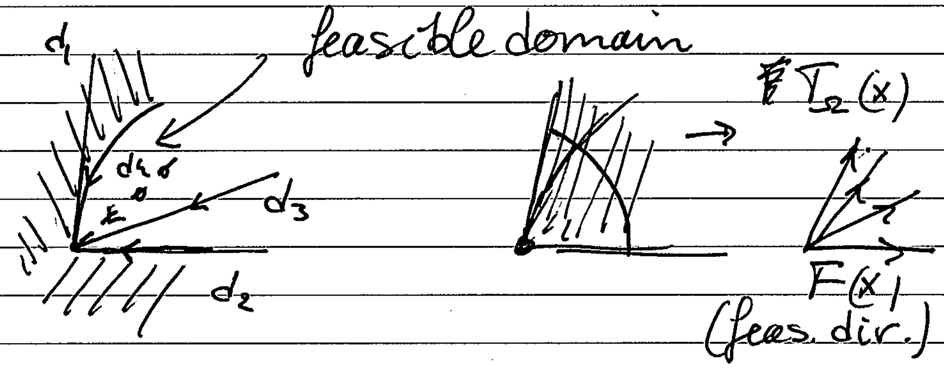
The set of all tangents to Ω at x is called the tangent cone and is denoted by $T_\Omega(x)$

Def 12.3

Given feas. point x and active constr. set $A(x)$, the set of linearized feas. directions $\mathcal{F}(x)$ is

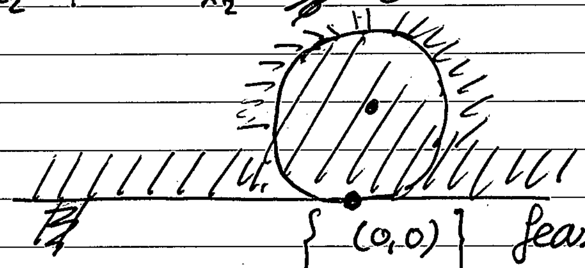
$$\mathcal{F}(x) = \left\{ d \;\middle|\; \begin{array}{ll} d^T \nabla c_i = 0 & \forall i \in \mathcal{E} \\ & \text{and} \\ d^T \nabla c_i \geq 0 & \forall i \in \mathcal{I} \cap \underbrace{A(x)}_{\text{active ineq. constr.}} \end{array} \right\}$$

(why compl. def? → tangent cone depends on geometry of Ω, not a choice of alg. specification.)



$(x_1-1)^2+(x_2-1)^2 \le 1$ ← $c_1 : 1-(x_1-1)^2-(x_2-1)^2 \ge 0$

$x_2 \le 0$ ← $c_2 : -x_2 \ge 0$



$\{(0,0)\}$ feasible set

$$\begin{cases} d^T \nabla c_1 = 0 \\ d^T \nabla c_2 = 0 \end{cases}$$

at $(0,0)$ true for $d = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$

but there is no direction to move

problem is that constr. are dependent.

Def 12.4 Given pt x and active set $A(x)$ the linear indep. constraint qualification (LICQ) holds if the set of active constraint grads $\{\nabla c_i \text{ for } i \in A(x)\}$ is linearly indep.

~~When~~

Relative to linearized constraints, what are necessary/suff. conditions of optimality?

Consider Lagrangian

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x)$$

First-order necessary conditions →

(nec. cond.s approx. to first order)

Theo ~~Def~~ 12.1

Suppose that x^* is a local solution of constr. min. problem in std form, that f and (all) c_i are cont. diff. and that LICQ holds at x^*. Then there is a Lagr. multiplier vector λ^* such that at (x^*, λ^*)

$$
\begin{aligned}
\nabla \mathcal{L}(x^*, \lambda^*) &= 0 \\
c_i(x^*) &= 0 \quad \forall i \in \mathcal{E} \\
c_i(x^*) &\geq 0 \quad \forall i \in \mathcal{I} \\
\lambda_i^* &\geq 0 \quad \forall i \in \mathcal{I} \\
{*1} \quad \lambda_i^* c_i(x^*) &= 0 \quad \forall i \in \mathcal{E} \cup \mathcal{I}
\end{aligned}
$$

Karush-Kuhn-Tucker conditions (KKT)

*1 complementarity conditions:

if $c_i(x^*) > 0$ then $\lambda_i = 0$ (constr. no effect)

if $c_i(x^*) = 0$ then $\lambda_i = 0$ or nonzer

↳ pref. $\lambda_i \neq 0$ otherwise some type of indeterminacy → active constr. does not influence Lagr. (locally)

Def 12.5

Strict Complementarity

Given local sol. x^* and λ^* satisfying ~~1st order nec~~ KKT cond.s, strict complementarity cond. holds if exactly one of $\lambda_i^* = 0$ or $c_i(x^*)$ holds for each $i \in \mathcal{I}$. In other words, $\lambda_i^* > 0$ for each $i \in \mathcal{I} \cap \mathcal{A}(x^*)$

When LICQ holds the (optimal) λ^* at local sol. x^* is unique

Def $A(x^*) = \left[\nabla c_i(x^*)^T \right]_{i \in \mathcal{A}(x^*)}$

Lemma: let x^* be feasible point. Then
(1) $T_\Omega(x) \subset F(x)$

(2) If LICQ satisfied at x, then $F(x) = T_\Omega(x)$

Proof: see book

Theo 12.3 (necessary cond)

If x^* is local solution, then

$$\nabla f(x^*)^T d \geq 0 \quad \text{for all } d \in T_\Omega(x^*)$$

Proof: Assume not true: $\exists d : \nabla f(x^*)^T d < 0$

Consider $\overset{\text{feas. seq.}}{\{z_k\}}, \{t_k\}$ s.t. $\lim_{k\to\infty} \frac{z_k - x^*}{t_k} = d$

$z_k = x^* + t_k d + o(t_k)$

$$f(z_k) = f(x^*) + (z_k - x^*)^T \nabla f(x^*) + o(\|z_k - x^*\|)$$

$$= f(x^*) + t_k d^T \nabla f^T(x^*) + o(t_k)$$

Since $d^T \nabla f^T(x^*) < 0$ (constant) and $t_k > 0$

for k large enough $f(z_k) < f(x^*)$

So, x^* not local solution (min in feas. set)

Lemma 12.4 (Farkas)

Let $B \in \mathbb{R}^{n \times m}$ and $C \in \mathbb{R}^{n \times p}$ and define cone $K = \{ By + Cw \mid y \geq 0 \}$. Then for any $g \in \mathbb{R}^n$ ~~either~~ exactly one of two cases holds

(either) (1) $g \in K$

comp. wise

(or) (2) $\exists d \in \mathbb{R}^n : g^T d < 0, B^T d \geq 0, C^T d = 0$

Assume f and c_i twice cont. diff.

Assume (x^*, λ^*) sat. KKT cond.s

Def. critical cone $C(x^*; \lambda^*) =$

$$\{w \in F(x^*) \mid \nabla c_i(x^*)^T w = 0 \text{ for all } i \in A(x^*) \cap I \text{ with } \lambda_i^* > 0\}$$

$$w \in C(x^*, \lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^T w = 0 & \forall i \in \mathcal{E} \\ \nabla c_i(x^*)^T w = 0 & \forall i \in A(x^*) \cap I \text{ w. } \lambda_i > 0 \\ \nabla c_i(x^*)^T w \geq 0 & \forall i \in A(x^*) \cap I \text{ w. } \lambda_i = 0 \end{cases}$$

$$\rightarrow \lambda_i^* w_i^T \nabla c_i(x^*) = 0 \quad \forall i \in \mathcal{E} \cup I$$

$$w \in C \Rightarrow w^T \nabla f(x^*) = \sum_{i \in \mathcal{E} \cup I} \lambda_i^* w_i^T \nabla c_i(x^*) = 0$$

↳ direction for which first derivative information does not give us info. whether f increases or decreases

Theo 12.5 Suppose x^* is local sol. and the LICQ is satisfied. Let λ^* be the Lagr. mult for which KKT satisfied. Then

$$w^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*) w \geq \text{ for all } w \in C(x^*, \lambda^*)$$

(second or necessary cond.)

Second order suff. cond.s

x^* feasible, KKT cond. hold

$$w^T \nabla_{xx} \mathcal{L}(x^*, \lambda^*) w > 0 \text{ for all } w \in C, \ w \neq 0$$

x^* is strict local minimizer