Ch 16: Quadratic Programming

* Subproblems to chapters 17-19.

QP: $\min_{x} \; q(x) = \frac{1}{2} x^T G x + x^T c$

subject to: $a_i^T x = b_i, \; i \in \mathcal{E}$

$a_i^T x \geq b_i, \; i \in \mathcal{I}$

G: pos. semidef is a <u>Convex QP</u>
(prob like LP)

16.1 Equality-Constr. QPs

$$\min_{x} \; q(x) \stackrel{\text{def}}{=} \frac{1}{2} x^T G x + x^T c$$

subject to: $Ax = b$.

Where A is $m \times n$ Jacobian of constr.
Assume A has full row-rank.

Soln:
$$\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix} \quad (*)$$

Set $x^* = x + p$. Then (*) becomes:

$$\underbrace{\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}}\begin{bmatrix} -p \\ \lambda^* \end{bmatrix} = \begin{bmatrix} g \\ h \end{bmatrix}$$

with: $h = Ax - b$, $g = c + Gx$, $p = x^* - x$.

KKT matrix is <u>always indefinite</u>!

The only way to attack this problem <u>directly</u> is to use special methods:

- <u>Symmetric indefinite factorization</u> (p455)
- <u>Schur-complement method</u> (p455)
- <u>Null-space method (p457)</u>

<u>Note:</u> Cholesky, CG will <u>not</u> work for singular matrices like this one.

- CG on nullspace method ok (p.459)

16.3 Iterative Solutions of the KKT System (for large problems)

- Algo 16.1 is a pre-conditiones CG for use with null-space matrix Z and Y so tha $[Y|Z]$ not singular (see p. 457)
- Algo 16.2 is the projected CG method.

16.4 Inequality-Constrained Problems

Define:

$$\mathcal{L}(x,\lambda) = \tfrac{1}{2}x^T G x + x^T c - \sum_{i \in I \cup \mathcal{E}} \lambda_i (a_i^T - b_i)$$

Recall the active set:

$$A(x^*) = \{ i \in \mathcal{E} \cup I \mid a_i^T x^* = b_i \}$$

KKT: $$Gx^* + c - \sum_{i \in A(x^*)} \lambda_i^* a_i = 0$$

$a_i^T x^* = b_i \; \overset{\text{for}}{\forall}\; i \in A(x^*)$, $a_i^T x^* \geq b_i$ for $i \in I - A(x^*)$

and $\lambda_i^* \geq 0$, all $i \in I \cap A(x^*)$.

<u>Thm</u> 16.4 says the solution is unique for G pos. def.

<u>Problems</u>:

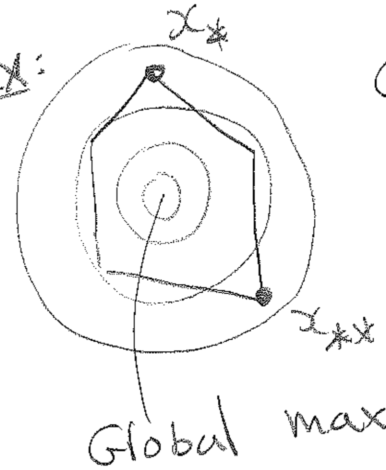
* a_i, $i \in A(x^*)$ are lin. dep. and/or
* Can have $\lambda_i^* = 0$ for some $i \in A(x^*)$

– Can cause zig-zags (p. 467)

* Non-convex problems

<u>Examples</u>:

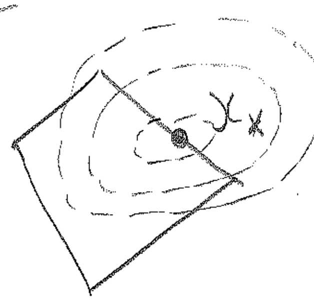
<u>Non-convex</u>:



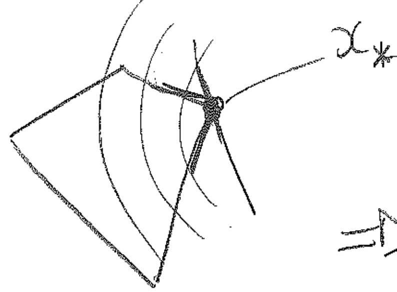
G with 2 negative eigenvalues

x_*, x_{**} are two local min

<u>Degenerate solution</u>



with $\nabla f(x_*) = 0$ (zero vector)



has <u>3</u> constraints in $\mathbb{R}^2$

$\Rightarrow$ Constraints are lin. dependent

Primal active set methods (p. 468)

Define a working set W_k at iteration k that includes all equality constraints and some of the inequality ones (ideally, pick the active ones).

Define: $x = x_k + p$, $g_k = Gx_k + c$

Subst in $q(x) = q(x_k + p) = \frac{1}{2} p^T G p + g_k^T p + \underbrace{\rho_k}_{\text{const w.r.t. } p}$

The problem becomes:

$$\min_p \ \frac{1}{2} p^T G p + g_k^T p$$

Subject to $a_i^T p = 0$, $i \in W_k$

which can be solved as in the direct methods or CG-based.

For line-search, consider

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

with $\alpha_k \overset{\text{def}}{=} \min\left(1, \min_{i \notin W_k} \frac{b_i - a_i^T x_k}{a_i^T p_k}\right)$

"Blocking constr."

Like LP: reducing directions $\rightsquigarrow a_i^T p_k < 0$

Note that if we <u>miss a constr</u>;

$$\frac{b_i - a_i^T x_k}{a_i^T \underline{p}_k} = \frac{0}{a_i^T \underline{p}_k} \text{ giving } \alpha_k = 0.$$

* This constr needs to be added to the working set.

* If $\alpha_k < 1$, then we have "blocking constraints". The new point will then activate this constr ($\Rightarrow$ going to zero) $\Rightarrow$ Update W_k by including it (or others if another dim is chosen) <u>Repeat</u>

At the solution, $\underline{p} = 0$ and all is satisfied. This gives:

$$\sum_{i \in \hat{W}} a_i \hat{\lambda}_i = g = G\hat{x} + c$$

for some multipliers $\hat{\lambda}_i$, $i \in \hat{W}$ and the rest are set to zero. Check for solution or iterate ...

Algorithm 16.3: Active-Set Method for Convex QP

Compute a feasible point x_0;

Set W_0 to be a subset of the active constraints at x_0;

(Set them to all "active" should work)

For $k = 0, 1, 2, \ldots$

Solve $\min_{p} \frac{1}{2} p^T G p + g_k^T p$

subject to $a_i^T p = 0, \ i \in W_k$

using a direct method or CG-based

If $p_k = 0$,

Solve $A^T \hat{\lambda} = g$ for $i \in W_k$

If $\hat{\lambda}_i \geq 0$ for $i \in W_k \cap I$

stop with $x^* = x_k$

else

$j \leftarrow \operatorname{argmin}_{j \in W_k \cap I} \hat{\lambda}_j$

$x_{k+1} \leftarrow x_k; \ W_k \leftarrow W_k \setminus \{j\}$

Remove constr with $\hat{\lambda}_j < 0$

else ($p_k \neq 0$)

$$\alpha_k = \min\left(1, \min_{i \notin W_k} \frac{b_i - a_i^T x_k}{a_i^T p_k}\right)$$

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

If ($\alpha_k < 1$ then there are blocking constraints.)

Obtain W_{k+1} by adding one of the blocking constr

(the one that gave min will work.)

else

$$W_{k+1} \leftarrow W_k$$

end (for)

Two methods are given on page 473 for how to pick x_0.

16.6 Interior-point Methods

A practical primal-dual method is given in algorithm 16.4 that computes iterates for (x, y, λ) where y are "slack variables".

Advice:

* For large problems, interior-point methods may lead to a solution faster
* When a good "warm start" x_0 is available, active-set methods may work very fast.

16.7 Gradient Projection Method

Problems of the form:

$$\min_x \; q(x) = \tfrac{1}{2} x^T G x + x^T c$$

$$\text{subject to: } l \leq x \leq u$$

are much easier to solve.

10/10

Algorithm 16.5 gives a fast method for solving this type of problem.

Note LBFGS solves this problem and you can just download the software!