Quadratic Programming

The general QP is

$$\min \ q(x) = \frac{1}{2} x^T G x + x^T c$$
$$\text{s.t.} \quad A_e x = b_e$$
$$A x \geq b$$
$$x \in \mathbb{R}^n$$

almost a linear program except for the quadratic term in the objective

or equivalently:

$$\min \ q(x) = \frac{1}{2} x^T G x + x^T c$$
$$\text{s.t.} \quad a_i^T x = b_i \quad i \in \mathcal{E}$$
$$a_i^T x \geq b_i \quad i \in \mathcal{I}$$
$$x \in \mathbb{R}^n$$

We will consider convex problems $(G \geq 0)$ which serve as subproblems for later methods.

Equality Constrained QP

$$\min \; q(x) = \frac{1}{2} x^T G x + x^T c$$
$$\text{s.t.} \quad Ax = b$$
$$x \in \mathbb{R}^n$$

The KKT conditions:

$$Gx + c - A^T \lambda = 0$$
$$Ax = b$$

which can be written in matrix form:

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

(First order necessary conditions)

Our assumptions:

(1) $G \geq 0$

(2) A is full rank

(3) $Z^T G Z > 0$

Example

min $\quad x^2 + y^2$

s.t. $\quad 3x + y = 3$

$$f(x) = x^2 + y^2 = \frac{1}{2} [x \ y] \overset{G}{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}} \begin{bmatrix} x \\ y \end{bmatrix} + [x \ y] \overset{c}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

$$3x + y = 3 \quad \Rightarrow \quad \overset{A}{[3 \ 1]} \begin{bmatrix} x \\ y \end{bmatrix} = \overset{b}{3}$$

KKT: $\begin{bmatrix} 2 & 0 & -3 \\ 0 & 2 & -1 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} x^* \\ y^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$

$$x_1^* = 0.9 \quad x_2^* = 0.3 \quad \lambda^* = 0.6$$
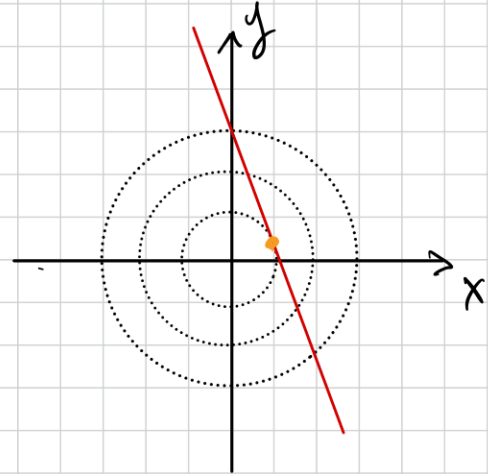
A computational variant

Suppose x not necessarily optimal.
Find a step p so that $x^* = x + p$.

$$\begin{cases} G(x+p) + c - A^T\lambda = 0 \\ A(x+p) = b \end{cases}$$

$$\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} p^* \\ x^* \end{bmatrix} = \begin{bmatrix} -Gx - c \\ b - Ax \end{bmatrix}$$

or

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -p^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} Gx + c \\ Ax - b \end{bmatrix}$$

Example: $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$Gx + c = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad Ax - b = 0$$

$$\begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \Rightarrow p^* = \begin{pmatrix} -0.1 \\ +0.3 \end{pmatrix}$$

$$x^* = x + p^* = \begin{pmatrix} 0.9 \\ 0.3 \end{pmatrix} \quad \checkmark$$

Example: $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$Gx + c = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad Ax - b = 1$$

$$\begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \Leftrightarrow p^* = \begin{pmatrix} -0.1 \\ -0.7 \end{pmatrix}$$

$$x^* = x + p^* = \begin{pmatrix} 0.9 \\ 0.3 \end{pmatrix} \quad \checkmark$$

An important condition

Let Z be the matrix whose columns form a basis for Null A. ($AZ=0$).

Lemma. Let A have full row rank and $Z^T G Z > 0$. Then
$$K = \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \text{ is invertible.}$$

Proof: Suppose w, v satisfying
$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \text{ If we}$$
show that $w=0$, $v=0$ is the unique solution then K is invertible.

First notice that $Aw = 0$. So,
$$0 = \begin{bmatrix} w^T & v^T \end{bmatrix} \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix}$$
$$= \begin{bmatrix} w^T & v^T \end{bmatrix} \begin{bmatrix} Gw + A^T v \\ Aw \end{bmatrix}$$
$$= w^T G w + w^T A^T v + v^T A w$$
$$= w^T G w$$

Now $w \in$ Null A so $\exists u \ni w = Zu$. So,
$$0 = w^T G w = (Zu)^T G (Zu) = u^T Z^T G Z u$$

Because $Z^T G Z > 0$, $u = 0$ and also $w = Zu = 0$.

Finally $Gw + A^T v = 0 \implies v = 0$. ▨

Theorem. Let A have full row rank and assume $Z^TGZ > 0$. Then x^* satisfying

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -x^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} c \\ -b \end{bmatrix} \quad \text{is the}$$
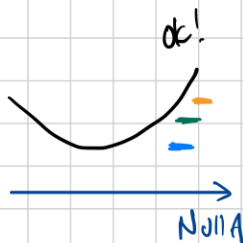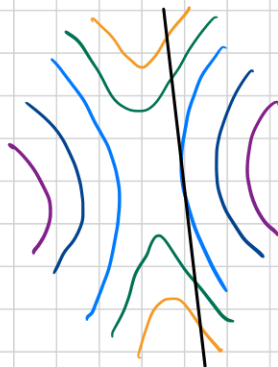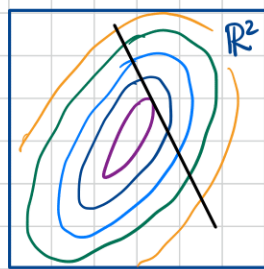
unique global minimizer of the ECQP.

Proof: Let $x \neq x^*$ be a feasible point and $p = x^* - x$. Using the facts that $Ap = 0$ and $p^T G x^* = p^T(-c + A^T \lambda^*) = -p^T c$, we have

$$q(x) = \tfrac{1}{2}(x^* - p)^T G (x^* - p) + c^T(x^* - p)$$
$$= q(x^*) + \tfrac{1}{2}p^T G p - p^T G x^* - c^T p$$
$$= q(x^*) + \tfrac{1}{2}p^T G p$$
$$> q(x^*).$$

Thus x^* is the unique global minimizer. ◻

Geometric interpretation

$\mathbb{R}^2$

ok!

Null A

ok!

Null A

Solving the KKT system

let $x^* = Y x_y + Z x_z$ (c)

Then $A(Z x_z) = (AZ) x_z = 0$

$$A(Y x_y) = b - AZ x_z = b$$

$$\min_x q(x) = \tfrac{1}{2} x^T G x + x^T c$$
$$\text{s.t.} \quad Ax = b$$

$$\min_{x_z} q(x) = \tfrac{1}{2} x_z^T Z^T G Z x_z + x_z^T (Z^T G Y x_y + Z^T c)$$

$$\Rightarrow \quad Z^T G Z x_z^* = -(Z^T G Y x_y^* + Z^T c) \quad \text{(a)}$$
$$\lessgtr \quad A Y x_y^* = b \quad \text{(b)}$$

Use (b) to solve for x_y^*.

Then use (a) to solve for x_z^*.

Then use (c) to find x^*.

The Projected Gradient CG Method

Basic Algorithm

Given: x satisfying $Ax = b$, $P = Z(Z^T H Z)^{-1} Z^T$

Compute: $r = Gx + c$, $g = Pr$, $d = -g$

Repeat:

$$\alpha \leftarrow r^T g / d^T G d$$

$$x \leftarrow x + \alpha d$$

$$r^+ \leftarrow r + \alpha G d$$

$$g^+ \leftarrow P r^+$$

$$\beta \leftarrow (r^+)^T g / r^T g$$

$$d \leftarrow -g^+ + \beta d$$

$$g \leftarrow g^+, \quad r \leftarrow r^+$$

Until $r^T g < tol$

✳ Does not use A or b!
The algorithm naturally operates in the affine subspace of the constraint set.

! Requires Null space basis matrix Z and preconditioner matrix H. These are, in general, not sparse.

A helpful modification

Choose $H = I$. Then

$$P = Z(Z^TZ)^{-1}Z^T = I - A^T(AA^+)^{-1}A$$

and

$$g = Pr = r - A^T(AA^T)^{-1}Ar$$
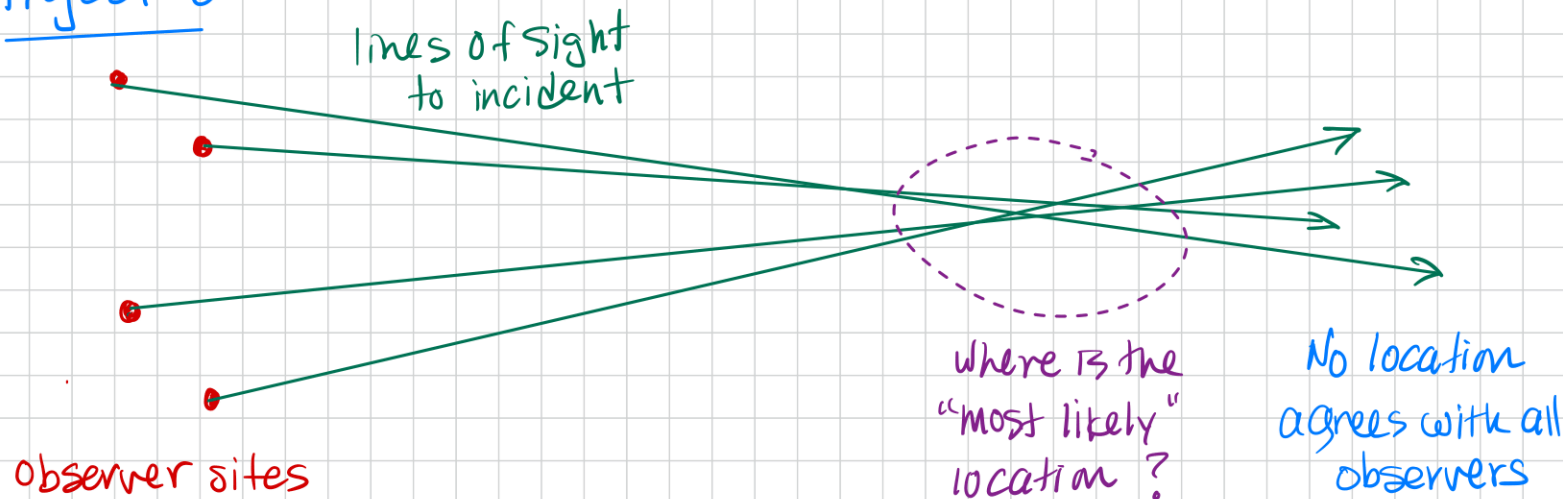$$g = Pr = r - A^Tv \quad , \quad AA^Tv = Ar$$

$\Rightarrow$ Procedure for computing g^+
    (a) solve $AA^Tv = Ar^+$ for v
    (b) $g^+ = r^+ - A^Tv$

— many variations exist —

Project 8

lines of sight
to incident

Observer sites

where is the
"most likely"
location?

No location
agrees with all
observers

we will formulate the problem in three different ways
and generalize to $\mathbb{R}^n$.

Method #1

A "solution" can be posed as the task of solving a system of linear equations, one for each line of sight.

$$Ax = b$$

which we can solve approximately using a pseudo-inverse via singular value decomposition (SVD).

$$A = U \Sigma V^T = \hat{U} \hat{\Sigma} \tilde{V}^T$$
$$P = \tilde{V} \hat{\Sigma}^{-1} \hat{U}^T$$
$$x^* = Pb$$

linear algebra

Method #2

Find the point that minimizes the sum of squared residuals:
For a point in $\mathbb{R}^n$ with m lines of sight:

$$\min \frac{1}{2} \|r\|^2$$
$$\text{s.t.} \quad r = Ax - b$$
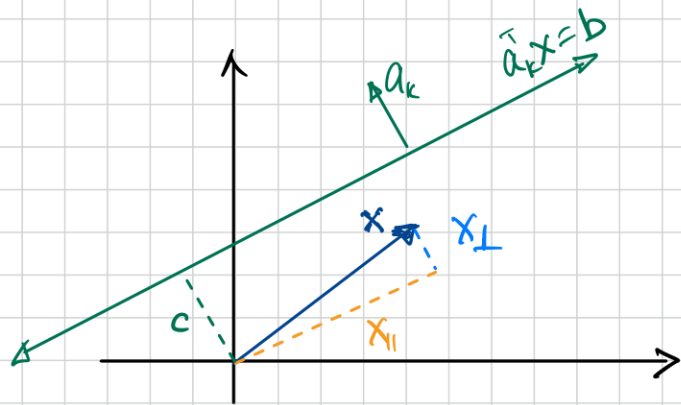$$x \in \mathbb{R}^n$$
$$r \in \mathbb{R}^m$$

quadratic program

Method #3

minimize the sum of squared distances to each line of sight.

The signed distance from a hyperplane defined by $a_k^T x = b_k$ to a point x is

$$d = \frac{b_k - a_k^T x}{\|a\|} = \frac{b_k}{\|a_k\|} - \frac{a_k^T x}{\|a_k\|}$$

$$x = x_\parallel + x_\perp$$

$$a^T x = a^T x_\parallel + a^T x_\perp$$

$$a^T x = 0 + \|a\| \|x_\perp\|$$

$$\|x_\perp\| = a^T x / \|a\|$$

when $a^T x = b$
then $\|x_\perp\| = c$

$$c = b / \|a\|$$

$$d = c - \|x_\perp\| = \frac{b}{\|a\|} - \frac{a^T x}{\|a\|}$$

So, one simple possibility is to normalize A and b and solve according to method #2.

$$\min \quad \frac{1}{2} \| r \|^2$$

$$\text{s.t.} \quad \| a_k \| r_k = a_k^T x - b_k$$

$$x \in \mathbb{R}^n$$

$$r \in \mathbb{R}^m$$

quadratic Program

Method #4

We can generalize the previous idea to find the point that minimizes total squared distances to a collection of polyhedral sets.
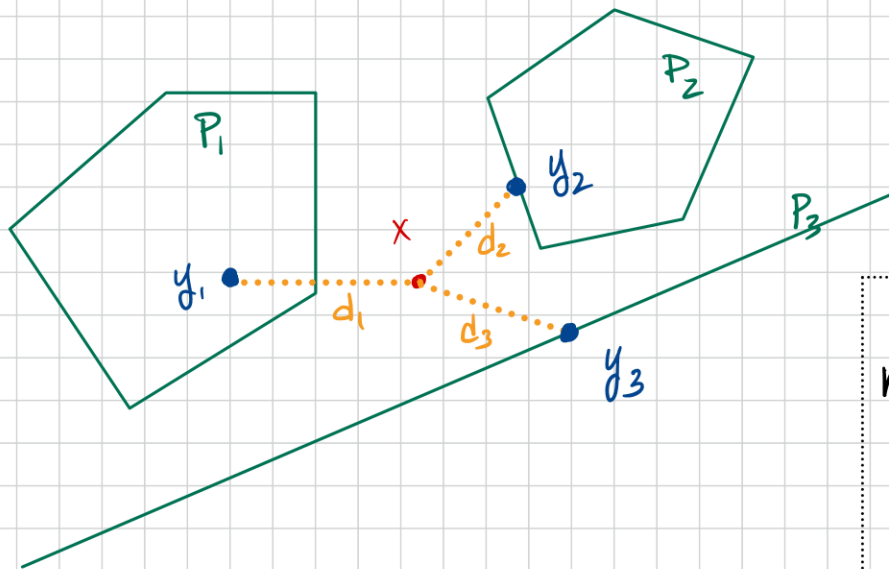
$$\min \quad \sum_{k=1}^{m} \| x - y_k \|^2 \quad + \alpha \| x \|^2$$

$$\text{s.t.} \quad y_k \in P_k$$

$$x \in \mathbb{R}^n$$

where $P_k = \{ x \in \mathbb{R}^n \mid A_k x = b_k, \ C_k x \geq d_k \}$

quadratic program

$P_1 = \{ x \in \mathbb{R}^n \mid A_1 x \leq b_1 \}$

$P_2 = \{ x \in \mathbb{R}^n \mid A_2 x \leq b_2 \}$

$P_3 = \{ x \in \mathbb{R}^n \mid A_3 x = b_3 \}$

$$\min \; \frac{1}{2} \sum_{k=1}^{3} \| x - y_k \|^2 + \frac{1}{2} \alpha \| x \|^2$$

s.t.
$A_1 y_1 \leq b_1$

$A_2 y_2 \leq b_2$

$A_3 y_3 = b_3$

$x, y_k \in \mathbb{R}^n \quad k = 1, 2, 3$

quadratic program with both inequality and equality constraints.

General Convex QP

Consider some of the first order necessary conditions:

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

$$\left. \begin{array}{ll} a_i^T x^* = b_i & i \in A(x^*) \\ a_i^T x^* \geq b_i & i \in I \setminus A(x^*) \\ \lambda_i^* \geq 0 & i \in I \cap A(x^*) \end{array} \right\} \circledast$$

Theorem. If x^* satisfies $\circledast$ for some λ_i^*, $i \in A(x^*)$, and $G \geq 0$, then x^* is a global minimizer of QP.

Proof: Suppose x is any other feasible point.

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

$$= \tfrac{1}{2} x^{*T} G x^* + x^{*T} c + (x-x^*)^T (G x^* + c) + \tfrac{1}{2}(x-x^*)^T G (x-x^*)$$

$$\geq q(x^*) + (x-x^*)^T (G x^* + c)$$

$$= q(x^*) + (x-x^*)^T \sum_{i \in A} \lambda_i^* a_i$$

$$= q(x^*) + \sum_{i \in A} \lambda_i^* a_i^T (x-x^*)$$

$$= q(x^*) + \sum_{i \in A \cap I} \lambda_i^* a_i^T (x-x^*)$$

$$= q(x^*) + \sum_{i \in A \cap I} \underset{\geq 0}{\lambda_i^*} (\underset{= b_i}{a_i^T x} - \underset{= b_i}{a_i^T x^*})$$
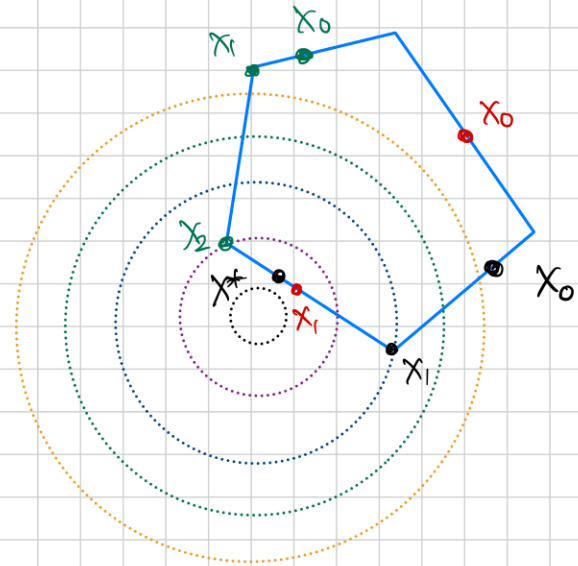
$$\geq q(x^*)$$

Active Set Method

(p 467-476)

If we know the active constraints at optimal point x^* the we could simply solve the equality constrained problem:

$$\min \quad q(x) = \frac{1}{2} x^T G x + x^T c$$

$$\text{s.t.} \quad a_i^T x = b_i \quad i \in A(x^*)$$

But we don't actually know $A(x^*)$.

Concept: Intelligently search combinations of active constraints (working set), improving $q(x)$ with a feasible sequence $x_k \Rightarrow x^*$.

First, reformulate the problem to get **step direction**

(i) Use a working set $W \subseteq A(x)$.

(ii) Use a minimizing KKT step P_k
on the working set constraints
$(x_{k+1} = x_k + \alpha P_k)$

let $p = x - x_k$
$g_k = G x_k + c$

Then
$q(x) = q(p + x_k)$

$= \frac{1}{2}(p + x_k)^T G (p + x_k) + (p + x_k)^T c$

$= \frac{1}{2} p^T G p + p^T G x_k + p^T c + \frac{1}{2} x_k^T G x_k + x_k^T c$

$-\frac{1}{2} p^T G p + g_k^T p + (\frac{1}{2} x_k^T G x_k + x_k^T c)$

$a_i^T p = a_i^T x - a_i^T x_k = b_i - b_i = 0 \quad \forall i \in W$

So, solve the following to find P_k:

$$\min_{P} \quad \frac{1}{2} p^T G p + g_k^T p$$
$$\text{s.t.} \quad a_i^T p = 0, \quad i \in W$$

The equality constrained
subproblem!

Constrained Newton step

Step Size

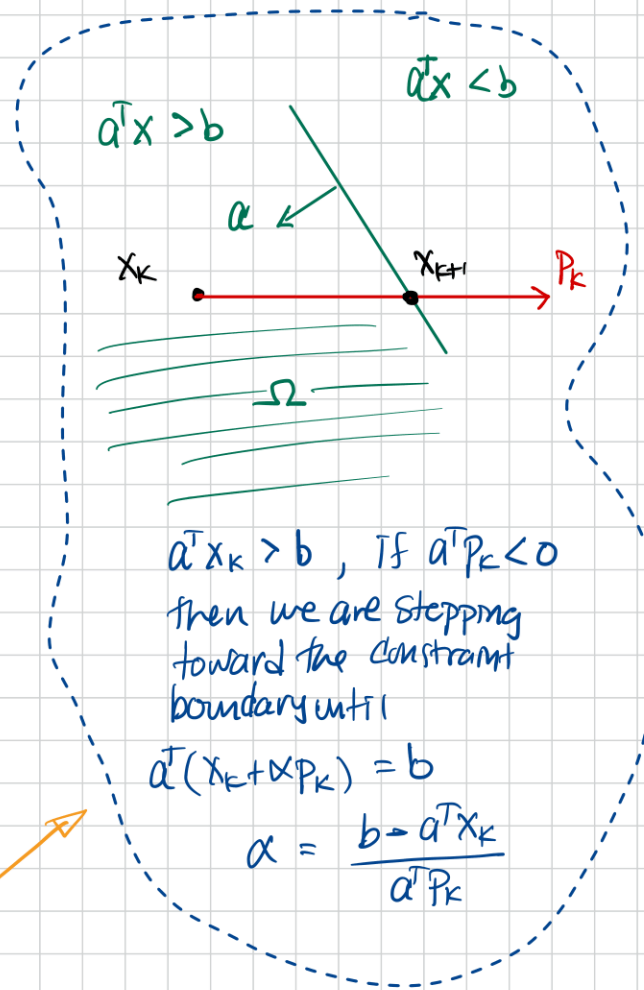$x_{k+1} = x_k + p_k$ may step to an infeasible point.

To guarantee feasibility, step only as far as possible in the direction of p_k.
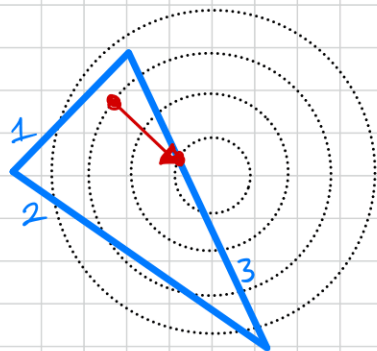
$$x_{k+1} = x_k + \alpha_k p_k$$

$$\alpha_k = \min\left\{ 1, \; \min_{\substack{a_i^T p_k < 0 \\ i \in W}} \frac{b_i - a_i^T x_k}{a_i^T p_k} \right\}$$
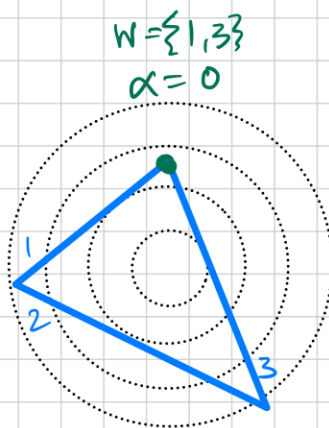
full step

Step to a "blocking constraint"

$a^T x > b$

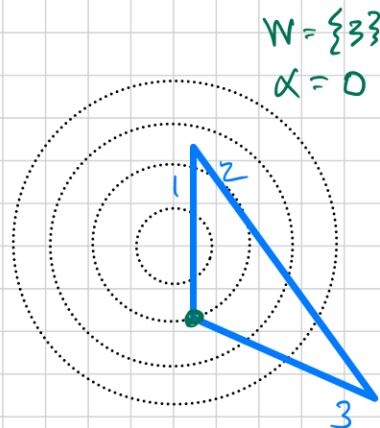$a^T x < b$

α

x_k x_{k+1} p_k

Ω

$a^T x_k > b$, if $a^T p_k < 0$ then we are stepping toward the constraint boundary until

$$a^T(x_k + \alpha p_k) = b$$

$$\alpha = \frac{b - a^T x_k}{a^T p_k}$$

Some Example Steps

$W = \phi$

$\alpha < 1$
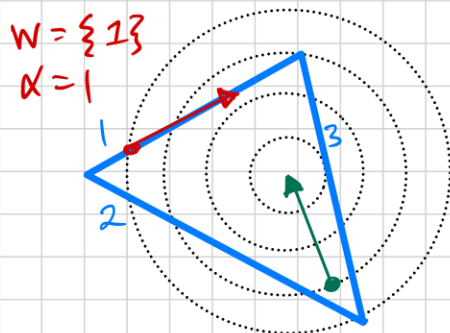
$W = \{1, 3\}$
$\alpha = 0$

$W = \{3\}$
$\alpha = 0$

$W = \{1\}$
$\alpha = 1$

$W = \phi$
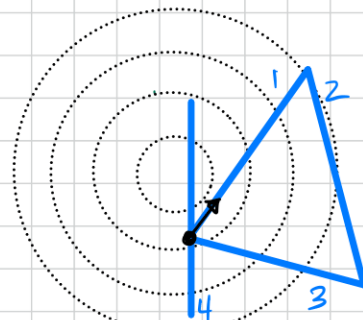$\alpha = 1$

$W = \{2, 3\}$
$\alpha = 0$

$W = \{3\}$
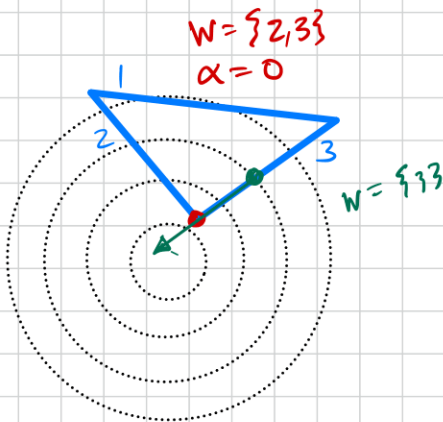
$W = \{4\}$ $W = \{1\}$
$\alpha = 0$ $\alpha = 1$

Algorithm Ideas

- If $\|P_k\| > 0$ and $x_k + P_k$ is feasible   (full Newton Step is feasible on W)

$$x_{k+1} = x_k + P_k$$

- If $\|P_k\| > 0$ and $x_k + P_k$ is infeasible   (Newton Step infeasible on W)

$$x_{k+1} = x_k + \alpha_k P_k \text{ and add the blocking constraint to W}$$

- If $\|P_k\| = 0$ and some $\lambda_\ell < 0, \ell \in W$ (a working set constraint prevents a step)

remove constraint ℓ from W

$$\sum_{i \in W} a_i \lambda_i = g = 6x + c$$

$$A_W \lambda_W = 6x + c$$

- If $\|P_k\| = 0$ and all $\lambda_\ell \geq 0, \ell \in W$ ($\vec{0}$ is the only feasible descent direction)

x_k is KKT optimal.

Active Set Algorithm

Algorithm 16.3 (Active-Set Method for Convex QP).

  Compute a feasible starting point x_0;

  Set $\mathcal{W}_0$ to be a subset of the active constraints at x_0;

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

      Solve (16.39) to find p_k;

      if $p_k = 0$

          Compute Lagrange multipliers $\hat{\lambda}_i$ that satisfy (16.42),

             with $\hat{\mathcal{W}} = \mathcal{W}_k$;

          if $\hat{\lambda}_i \geq 0$ for all $i \in \mathcal{W}_k \cap \mathcal{I}$

             stop with solution $x^* = x_k$;

          else

             $j \leftarrow \arg\min_{j \in \mathcal{W}_k \cap \mathcal{I}} \hat{\lambda}_j$;

             $x_{k+1} \leftarrow x_k$; $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k \backslash \{j\}$;

      else (* $p_k \neq 0$ *)

          Compute α_k from (16.41);

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

          if there are blocking constraints

             Obtain $\mathcal{W}_{k+1}$ by adding one of the blocking

                constraints to $\mathcal{W}_k$;

          else

             $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k$;

  end (for)

Let $\mathcal{W}_0$ be the set of equality constraints, which are assumed to be linearly independent.

Finding an Initial Feasible Point

Suppose we have an initial guess $\hat{x}$ which is not feasible.

Let $s_i = \text{sign}(a_i^T \hat{x} - b_i)$, $i \in \mathcal{E}$.

Then we solve the linear program

$$
\begin{aligned}
\min_{(x,y)} \quad & \sum_{i \in \mathcal{E} \cup \mathcal{I}} y_i \\
\text{s.t.} \quad & a_i^T x - s_i y_i = b_i, \quad i \in \mathcal{E} \\
& a_i^T x + y_i \geq b_i, \quad i \in \mathcal{I} \\
& y \geq 0
\end{aligned}
$$

If $y^* = 0$ then x^* is a feasible point for the QP.

And the active set $A(x^*, y^*)$ for $i \in \mathcal{E} \cup \mathcal{I}$ or $i \in \mathcal{E}$ can be used as an initial working set for the QP.