

# Quadratic Programming

The general QP is

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

$$\text{s.t. } A_e x = b_e$$

$$A x \geq b$$

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

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

or equivalently:

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

$$\text{s.t. } a_i^T x = b_i \quad i \in \mathcal{E}$$

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

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

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

## Equality Constrained QP

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

$$\text{s.t. } Ax = b$$

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

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)

The KKT conditions:

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

$$Ax = b$$

Our assumptions:

(1)  $G \geq 0$

(2)  $A$  is full rank

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

## Example

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

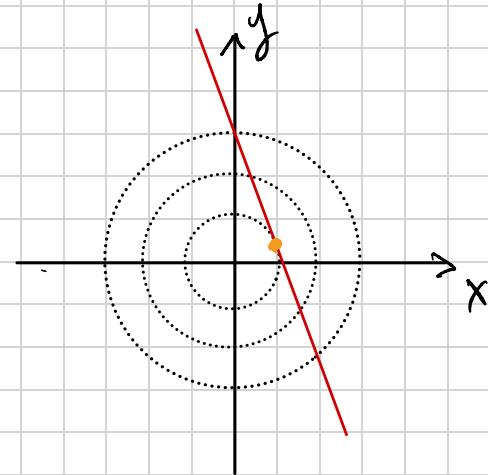
$$\text{s.t.} \quad 3x + y = 3$$

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

$$3x + y = 3 \Rightarrow \begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}$$

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

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



## A computational Variant

Suppose  $x$  not necessarily optimal.

Find a step  $p$  so that  $x^* = x + p$ .

$$\left\{ \begin{array}{l} G(x+p) + c - A^T \lambda = 0 \\ A(x+p) = b \end{array} \right.$$



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

or

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

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

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

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

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

## An important condition

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

**Lemma.** Let  $A$  have full row rank and  $Z^T G Z > 0$ . Then

$$K = \begin{bmatrix} G & A^T \\ A & O \end{bmatrix} \text{ is invertible.}$$

**Proof:** Suppose  $w, v$  satisfying

$$\begin{bmatrix} G & A^T \\ A & O \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \text{ If we}$$

Show that  $w=0, v=0$  is the unique solution then  $K$  is invertible.

First notice that  $Aw=0$ . So,

$$0 = [w^T \ v^T] \begin{bmatrix} G & A^T \\ A & O \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix}$$

$$= [w^T \ v^T] \begin{bmatrix} Gw + A^T v \\ Aw \end{bmatrix}$$

$$= w^T Gw + w^T A^T v + v^T Aw$$

$$= w^T Gw$$

Now  $w \in \text{Null } A$  so  $\exists u \ni w = zu$ , so,

$$0 = w^T Gw = (zu)^T G (zu) = u^T z^T G Z u$$

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

Finally  $Gw + A^T v = 0 \Rightarrow v=0$ . \(\blacksquare\)

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

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

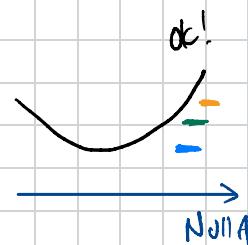
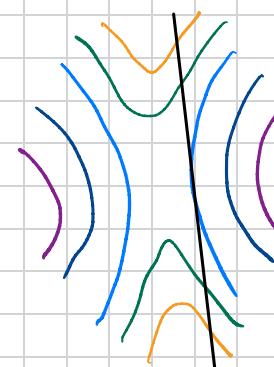
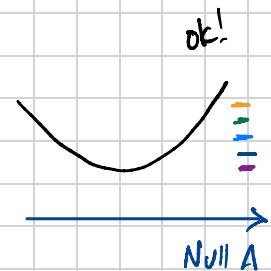
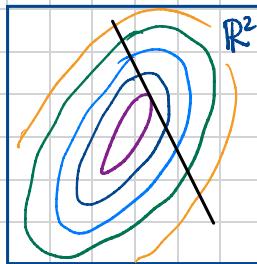
unique global minimizer of the ECQP.

**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

$$\begin{aligned} q(x) &= \frac{1}{2} (x^* - p)^T G (x^* - p) + c^T (x^* - p) \\ &= q(x^*) + \frac{1}{2} p^T G p - p^T G x^* - c^T p \\ &= q(x^*) + \frac{1}{2} p^T G p \\ &> q(x^*). \end{aligned}$$

Thus  $x^*$  is the unique global minimizer.  $\square$

Geometric interpretation



## Solving the KKT system

$$\text{let } X^* = YX_y + ZX_z \quad (c)$$

$$\text{Then } A(ZX_z) = (AZ)X_z = 0$$

$$A(YX_y) = b - AZX_z = b$$

$$\begin{aligned} \min_{\mathbf{x}} \quad & q(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T G \mathbf{x} + \mathbf{x}^T C \\ \text{s.t.} \quad & A \mathbf{x} = b \end{aligned} \quad \left. \right\} \rightarrow$$

$$\min_{X_z} \quad q(X) = \frac{1}{2} X_z^T Z^T G Z X_z + X_z^T (Z^T G Y X_y + Z^T C)$$

$$\Rightarrow \boxed{Z^T G Z X_z^* = -(Z^T G Y X_y^* + Z^T C)} \quad (a)$$

$\Leftrightarrow A Y X_y^* = b$  (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:  $r = b - Ax$ ,  $g = Pr$ ,  $d = -g$

Repeat:

$$x \leftarrow r^T g / d^T b d$$

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

$$r^+ \leftarrow r + \alpha Gd$$

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

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

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

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

Until  $r^T g < tol$

! Does not use  $A$  or  $b$ !

The algorithm naturally operates in the affine subspace of the constraint set.

! Requires  $M \times N$  space basis matrix  $Z$  and preconditioner matrix  $H$ . These are, in general, not sparse.

## A helpful modification

Choose  $H = I$ . Then

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

and

$$g = Pr = r - A^T (A A^T)^{-1} A r$$

$$g = Pr = r - A^T v, \quad A A^T v = A r$$

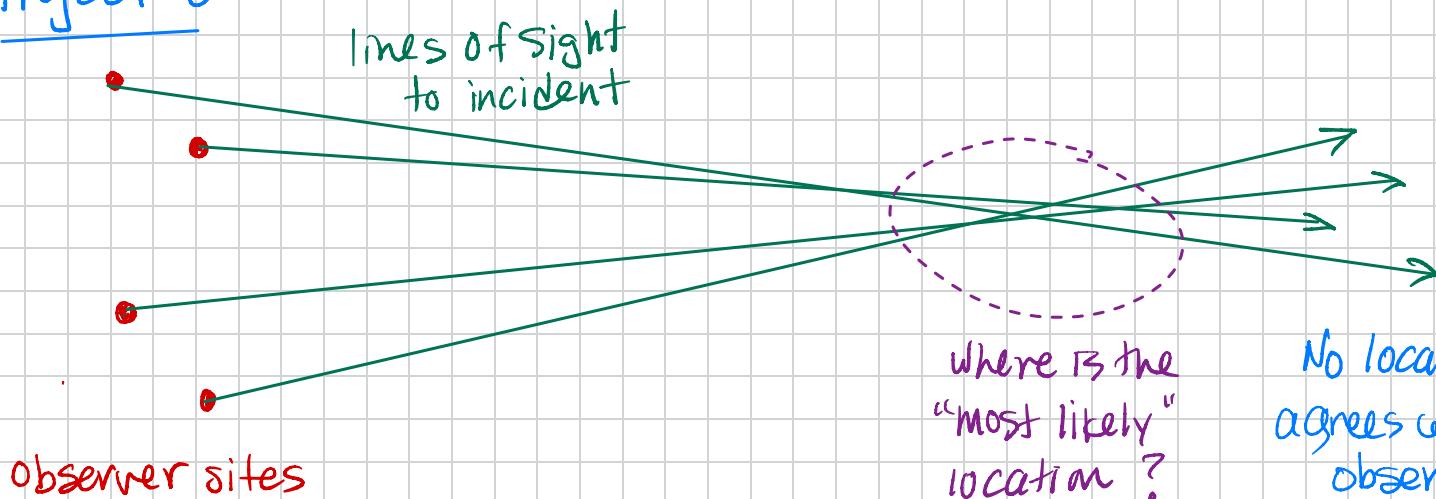
— Many variations exist —

⇒ Procedure for computing  $g^+$

(a) solve  $A A^T v = A r^+$  for  $v$

(b)  $g^+ = r^+ - A^T v$

## Project 8



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).

$$A = U \Sigma V^T = \tilde{U} \tilde{\Sigma} \tilde{V}^T$$

$$P = \tilde{V} \tilde{\Sigma}^{-1} \tilde{U}^T$$

$$x^* = Pb$$

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:

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

$$\text{s.t. } r = Ax - b$$

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

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

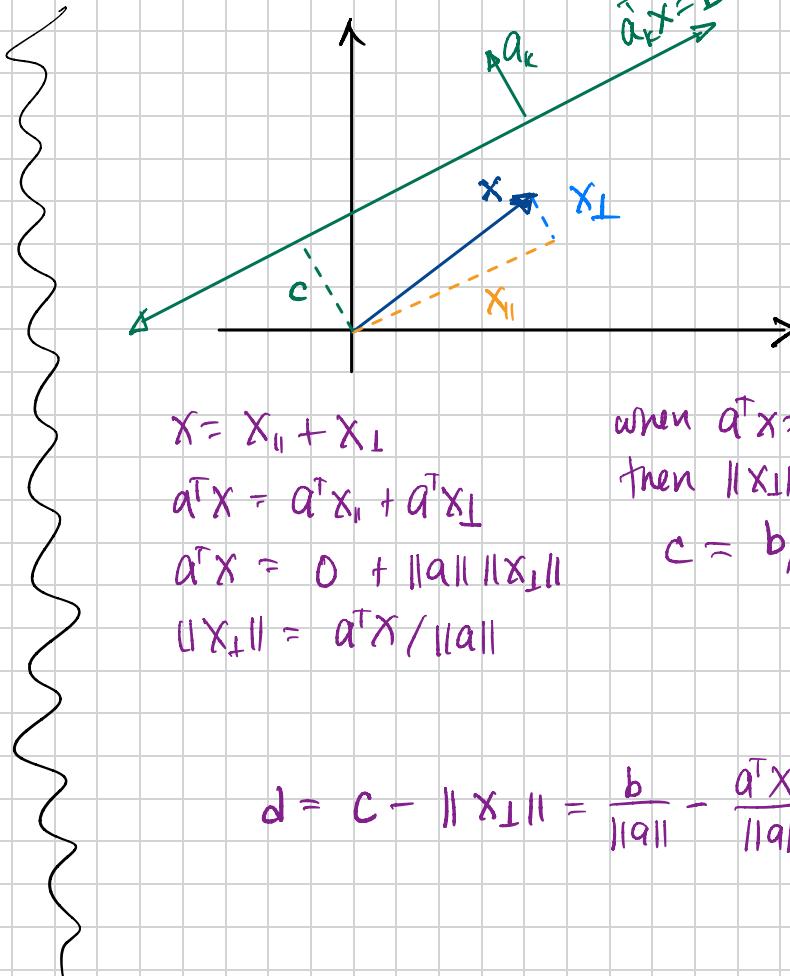
quadratic  
program

### 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

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



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

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

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

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

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.

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

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

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

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

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.

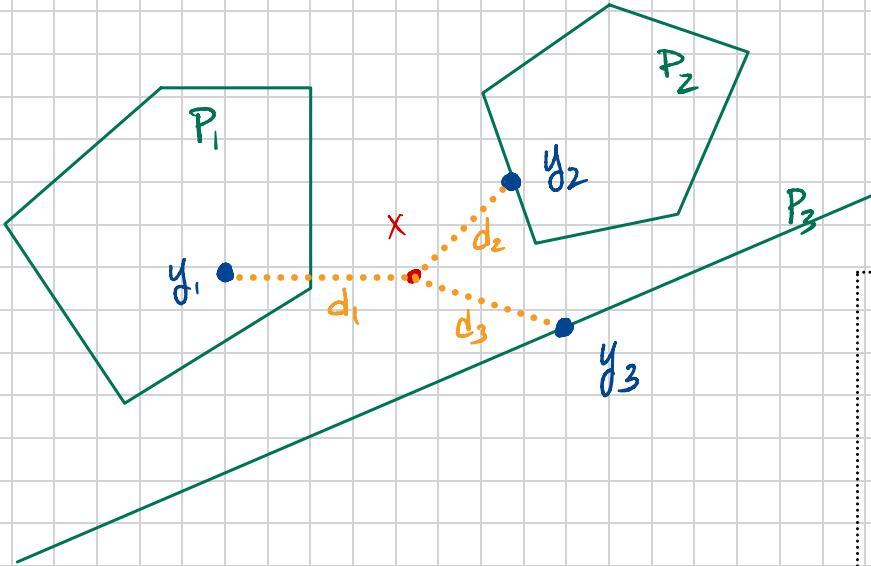
$$\min \sum_{k=1}^m \|x - y_k\|^2 + \alpha \|x\|^2$$

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

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

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

quadratic program



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

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

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

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

$$\text{s.t. } A_1 y_1 \leq b_1$$

$$A_2 y_2 \leq b_2$$

$$A_3 y_3 = b_3$$

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

quadratic program with  
both inequality and  
equality constraints.

## General Convex QP

Consider some of the first order necessary conditions:

$$\begin{aligned} Gx^* + c - \sum_{i \in A(x^*)} \lambda_i^* a_i &= 0 \\ a_i^T x^* = b_i &\quad i \in A(x^*) \\ a_i^T x^* \geq b_i &\quad i \in I \setminus A(x^*) \\ \lambda_i^* \geq 0 &\quad i \in I \cap A(x^*) \end{aligned}$$

(\*)

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

**Proof:** Suppose  $x$  is any other feasible point.

$$\begin{aligned} q(x) &= \frac{1}{2} x^T G x + c^T x \\ &= \frac{1}{2} x^{*T} G x^* + x^{*T} c + (x-x^*)^T (Gx^* + c) + \frac{1}{2} (x-x^*)^T G (x-x^*) \\ &\geq q(x^*) + (x-x^*)^T (Gx^* + c) \\ &= q(x^*) + (x-x^*)^T \sum_{i \in A} \lambda_i^* a_i \\ &= q(x^*) + \sum_{i \in A} \lambda_i^* a_i^T (x-x^*) \\ &= q(x^*) + \sum_{i \in A \cap I} \lambda_i^* a_i^T (x-x^*) \\ &= q(x^*) + \sum_{i \in A \cap I} \lambda_i^* (a_i^T x - a_i^T x^*) \\ &\geq q(x^*) \end{aligned}$$

□ □ □



## Active Set Method

(p 467-476)

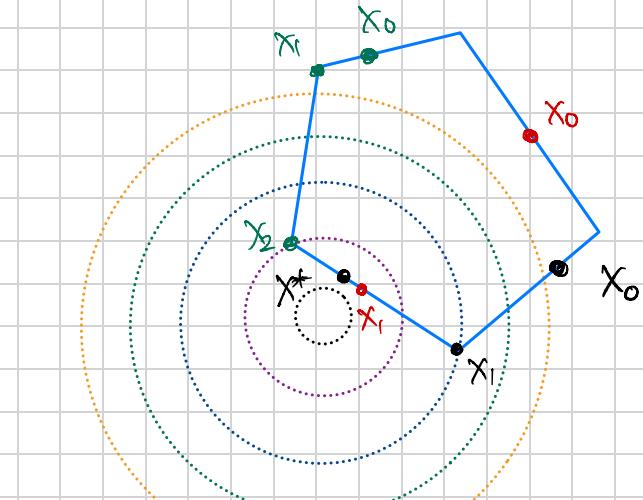
If we know the active constraints at optimal point  $x^*$  then we could simply solve the equality constrained problem:

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

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

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  $P = x - x_k$

$$g_k = Gx_k + c$$

Then

$$g(x) = g(P + x_k)$$

$$= \frac{1}{2}(P + x_k)^T G(P + x_k) + (P + x_k)^T C$$

$$= \frac{1}{2}P^T GP + P^T Gx_k + P^T C + \frac{1}{2}x_k^T Gx_k + x_k^T C$$

$$- \frac{1}{2}P^T GP + g_k^T P + \left( \frac{1}{2}x_k^T Gx_k + x_k^T C \right)$$

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

So, solve the following to find  $P_k$ :

$$\begin{aligned} & \min_P \frac{1}{2} P^T GP + g_k^T P \\ & \text{s.t. } a_i^T P = 0, i \in W \end{aligned}$$

The equality constrained subproblem!

Constrained Newton step

## Step Size

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

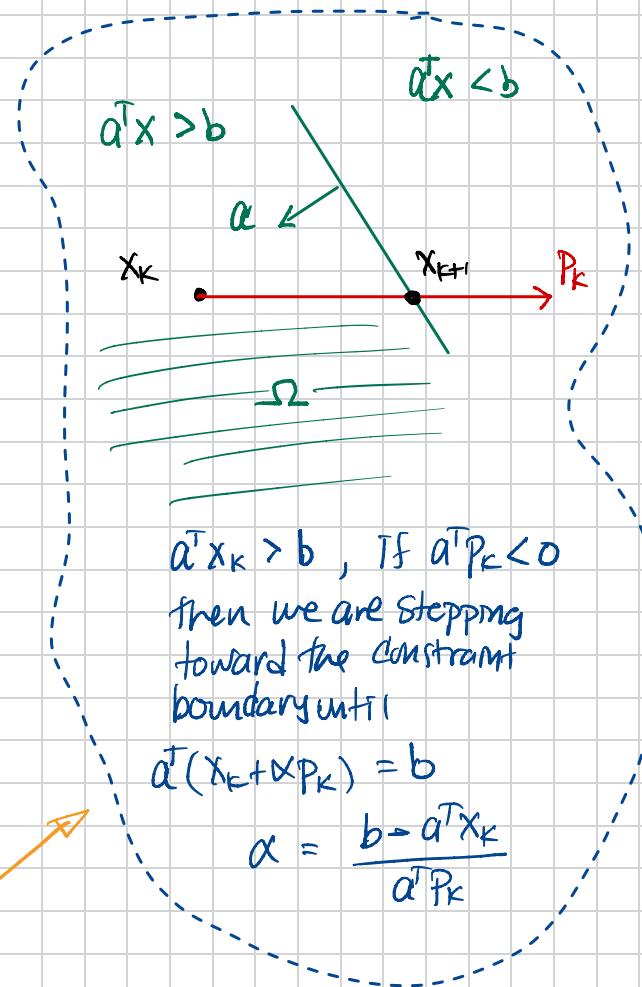
To guarantee feasibility, step only as far as possible in the direction of  $p_k$ .

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

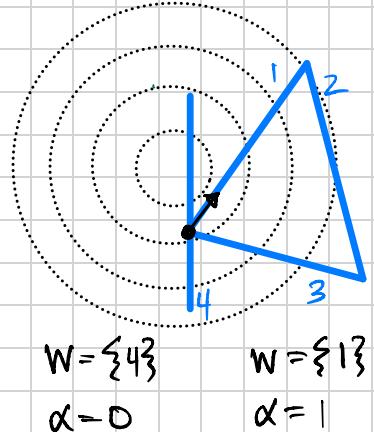
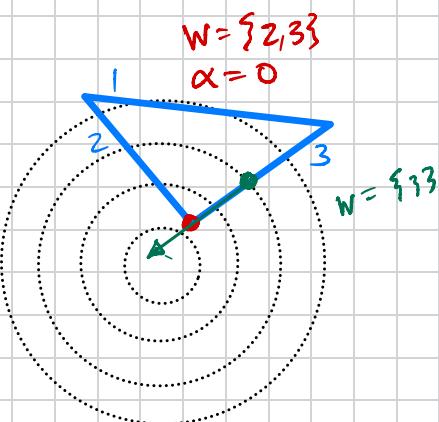
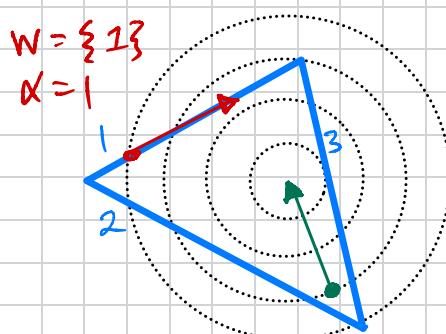
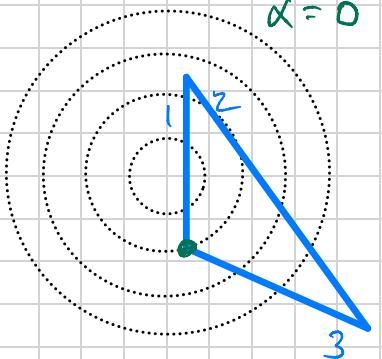
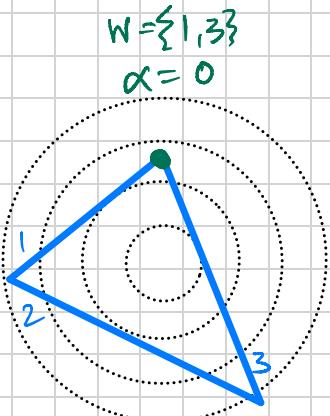
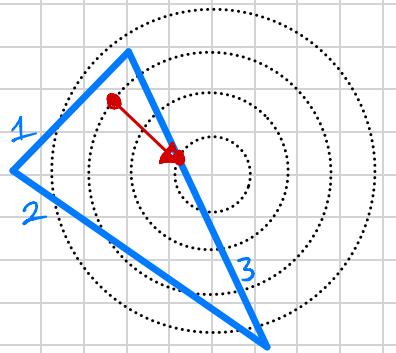
$$\alpha_k = \min \left\{ 1, \min_{\substack{i \in \mathcal{W} \\ a_i^T p_k < 0}} \frac{b_i - a_i^T x_k}{a_i^T p_k} \right\}$$

full step

Step to a "blocking constraint"



## Some Example Steps



## 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  $W$ )  
$$x_{k+1} = x_k + \alpha_k P_k$$
 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  $\ell$  from  $W$   
$$\sum_{i \in W} a_i \lambda_i = g = Gx + C$$
$$A_W \lambda_W = Gx + C$$
- 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**  $k = 0, 1, 2, \dots$

    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**

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

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

**else** (\*  $p_k \neq 0$  \*)

            Compute  $\alpha_k$  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.

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

Then we solve the linear program

$$\begin{array}{ll}\text{Min}_{(x,y)} & \sum_{i \in \Sigma \cup \mathcal{E}} y_i\end{array}$$

$$\text{s.t. } a_i^T x - s_i y_i = b_i, i \in \mathcal{E}$$

$$a_i^T x + y_i \geq b_i, i \in \Sigma$$

$$y \geq 0$$

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

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

