QUADRATIC PROGRAMMING BASICS

H. E. Krogstad

Spring 2005/Rev. 2008

***Quadratic Programming* (QP):**

- Common form for a lot of problems
- The iteration step in *Sequential Quadratic Programming* (SQP) methods

THE QP PROBLEM

We are considering problems where the objective function is *quadratic*,

$$q(x) = \frac{1}{2}x'Gx + d'x, \quad G \text{ symmetric.}$$

For the non-constrained problem we know

$$\nabla q(x)' = Gx + d,$$

$$x^* = -G^{-1}d \text{ when } G > 0$$

In general, *G will not necessarily be positive definite, not even semi-definite.*

The *feasibility domain* Ω is defined in terms of

- **linear *equality constraints,***

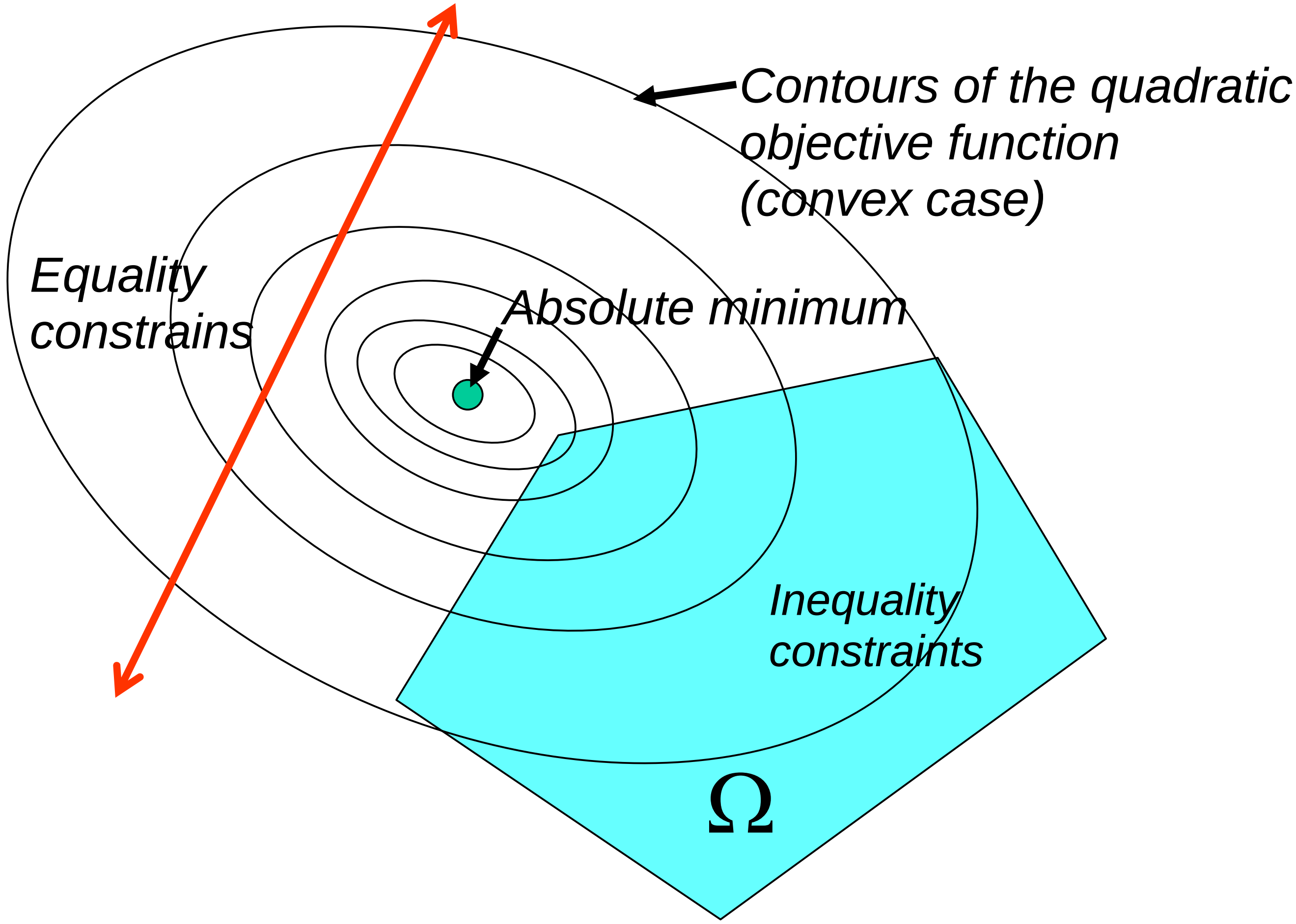
$$a_i' x = b_i, \; i \in \mathcal{E},$$

- **linear *inequality constraints,***

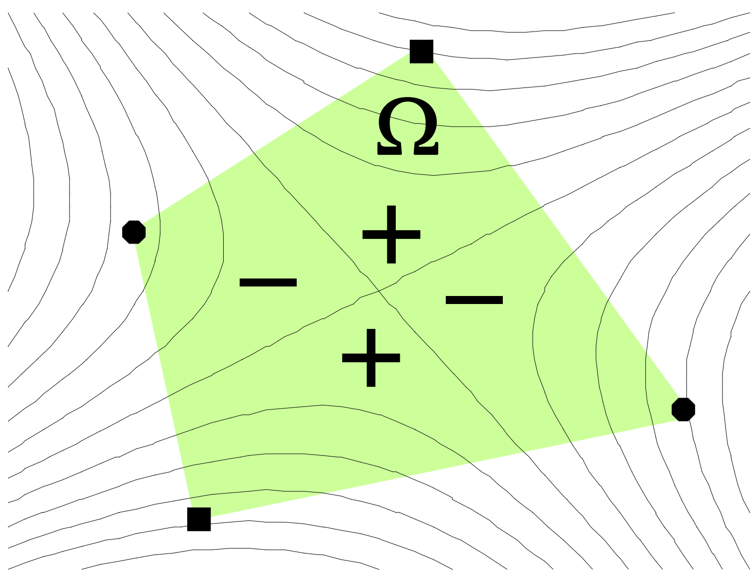
$$a_i' x \geq b_i, \; i \in \mathcal{I}.$$

NOTE:

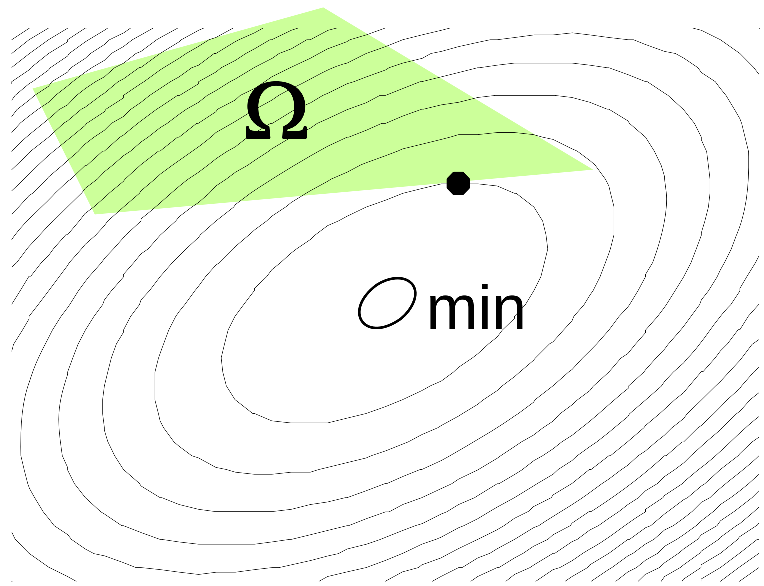
- Ω is convex
- The objective function will be convex if $G \geq 0$.


Contours of the quadratic objective function (convex case)
Equality constrains
Absolute minimum
Inequality constraints
Ω

Indefinite



Positive definite



An indefinite matrix G may lead to several local minima!

THE QP PROBLEM WITH EQUALITY CONSTRAINTS

This case may always be reduced to the following:

$$\min q(x)$$

$$Ax = b,$$

$$A \text{ has full rank } r < n.$$

$$\Omega = \{x;\ Ax = b\} = \{x_0 + Zu;\ Ax_0 = b,\ Z \text{ contains a basis for } N(A),\ u \in R^{n-r}\}.$$

Solution by the Null-Space Method:

Insert $x = x_0 + Zu$:

$$f(u) = q(x_0 + Zu) = \frac{1}{2}(x_0 + Zu)' G(x_0 + Zu) + d'(x_0 + Zu)$$

$$= \frac{1}{2}u'\tilde{G}u + \tilde{d}'u + \text{const.}, \quad \tilde{G} = Z'GZ, \ \ \tilde{d} = Z'(Gx_0 + d).$$

$$\nabla f(u) = \tilde{G}u + \tilde{d} = 0.$$

Three cases, depending on $\tilde{G}$:

1) A unique solution if the matrix is positive definite

2) Infinitely many solutions if $\tilde{G}$ is singular, as long as it is positive semi-definite and

$$\tilde{d} \in \mathcal{R}(\tilde{G}) .$$

3) No solutions if it is not positive semi-definite (or $\tilde{d} \notin R(\tilde{G})$)

Solving the KKT-equations directly

The Lagrange function is

$$L(x,\ \lambda) = \frac{1}{2}x'Gx + d'x - \lambda'(b - Ax),$$

and hence,

$$\nabla_x L(x, \lambda)' = Gx + d + A'\lambda = 0,$$
$$Ax = b.$$

Collected into a system:

$$\begin{bmatrix} G & A' \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -d \\ b \end{bmatrix}.$$

Lemma 16.1: *The coefficient matrix of the system is non-singular if A has full rank and G is positive definite on the null-space of A.*

Assume that A has full rank and G is positive definite on the null space of A, that is

$$(Zu)'G(Zu) = u'Z'GZu = u'\tilde{G}u > 0 \ \forall u \neq 0.$$

Then (x^*,λ^*) is a unique KKT point and a global minimum

(Follows from the Null-Space Method, which then solves a strictly convex problem).

There are many ways of solving the KKT system in this case. However, if x^* is known,

$$\lambda^* = -(AA')^{-1} A(Gx^* + d)$$

If λ^* is known, we reduce the over-determined system

$$Gx = -(d + A'\lambda^*)$$
$$Ax = b$$

to a (non-singular) system with n unknowns (Simple if $G > 0$!)

INEQUALITY CONSTRAINTS

Inequality constraints spoil the elegant theory above completely!

Let us consider the general problem

$$\min_x \left\{ \frac{1}{2} x' G x + d' x \right\},$$

$$a_i' x_i - b = 0, \; i \in \mathrm{E},$$

$$a_i' x - b_i \geq 0, \; i \in \mathrm{I}.$$

We assume that all *equality constraints* are linearly independent.

Let as $\mathcal{A}$ or $\mathcal{A}(x)$ denote the active set of constraints in x

The KKT-conditions:

$$
\begin{aligned}
Gx + d - \sum_{i \in \mathbf{A}(x)} \lambda_i a_i &= 0, \\
a_i' x = b_i, &\quad i \in \mathsf{A}(x), \\
a_i' x > b_i, &\quad i \in \mathtt{I} \setminus \mathsf{A}(x), \\
\lambda_i \geq 0, &\quad i \in \mathtt{I} \cap \mathsf{A}(x)
\end{aligned}
$$

(Recall that the LICQ conditions are *not necessary* for linear constraints).

If x^* is a KKT-point for the full problem, then x^* and a corresponding subset of the Lagrange multipliers is also a KKT-point for the ***reduced problem***:

$$
\min_x \left\{ \frac{1}{2} x' G x + d' x \right\},
$$

$$
a_i' x = b_i, \ i \in \mathcal{A}(x^*)
$$

- The reduced problem is a QP problem with equality constraints.

- If we have an active set $\mathcal{A}$ and have found a KKT-point x^* for the reduced problem, it is easy to check the KKT-conditions for the full problem.

- The next step would be to check the 2nd order conditions: Form the matrix A consisting of the gradients of the active constraints and investigate $Z'GZ$, where Z is a basis of $\mathcal{N}(A)$.

ACTIVE SET METHODS

1. Assume we are in a point $x_0 \in \Omega$.

2. We choose a *working set* $\mathcal{W}$ so that

$$\mathcal{E} \subset \mathcal{W} \subset \mathcal{A}(x_0).$$

3. Let A_W be the corresponding matrix of gradients and solve equality constrained reduced problem

$$\min q(x_0 + p),$$
$$A_W(x_0 + p) = b_W.$$

4. If p turns out to be 0 we have to check whether x_0 could be the full solution.

5. If $p \neq 0$, we consider p as a search direction and determine $\alpha \leq 1$ as the maximum value where

$$x_1 = x_0 + \alpha p \in \Omega.$$

6a. If $\alpha = 1$, we are at a KKT-point for the reduced problem ($x_1 \in \Omega$!).

6b. Otherwise, new inequality constraints have become active, which we now include in $\mathcal{W}$, and continue as above from x_1.

When this stops, we have a point x^* and an active set $\mathcal{W}^*$.

This point satisfies

$$Gx^* + d - \sum_{i \in \mathcal{W}^*} \lambda_i^* a_i = 0.$$

7a. Set the Lagrange multipliers for the constraints that are not in $\mathcal{W}^*$ to be 0.

If

$$\lambda_i^* \geq 0$$

for all $i \in \mathcal{W}^* \cap \mathcal{I}$, we have reached a KKT-point for the full solution and that needs to be checked (unless we have a convex problem).

7b. However, if some of the these multipliers are negative, we throw the corresponding constraints out from $\mathcal{W}^*$ and solve a new reduced problem.

(It may be shown, Theorems 16.5 and 16.6 in N&W, that this will decrease the objective further!).

Note:

- In order to start the method, that is to identify a feasible point $x_0 \in \Omega$, it may be necessary to carry out a Phase 1 problem as in the LP case.

- The Active Set algorithm is listed on p. 472, and numerical aspects are given on p. 477-480.

Algorithm 16.1 (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 (16.27) to find p_k;
    if $p_k = 0$
        Compute Lagrange multipliers $\hat{\lambda}_i$ that satisfy (16.30),
            set $\hat{\mathcal{W}} = \mathcal{W}_k$;
        if $\hat{\lambda}_i \geq 0$ for all $i \in \mathcal{W}_k \cap \mathcal{I}$;
            STOP with solution $x^* = x_k$;
        else
            Set $j = \arg\min_{j \in \mathcal{W}_k \cap \mathcal{I}} \hat{\lambda}_j$;
            $x_{k+1} = x_k$; $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k \backslash \{j\}$;
    else (* $p_k \neq 0$ *)
        Compute α_k from (16.29);
        $x_{k+1} \leftarrow x_k + \alpha_k p_k$;
        if there are blocking constraints
            Obtain $\mathcal{W}_{k+1}$ by adding one of the blocking
                constraints to $\mathcal{W}_{k+1}$;
        else
            $\mathcal{W}_{k+1} \leftarrow \mathcal{W}_k$;
end (for)

(Copied from Ed. 1. See p.462, 2nd Ed.)

Example 16.3:

$$\min_x \left\{ (x_1 - 1)^2 + \left(x_2 - \frac{5}{2} \right)^2 \right\}$$

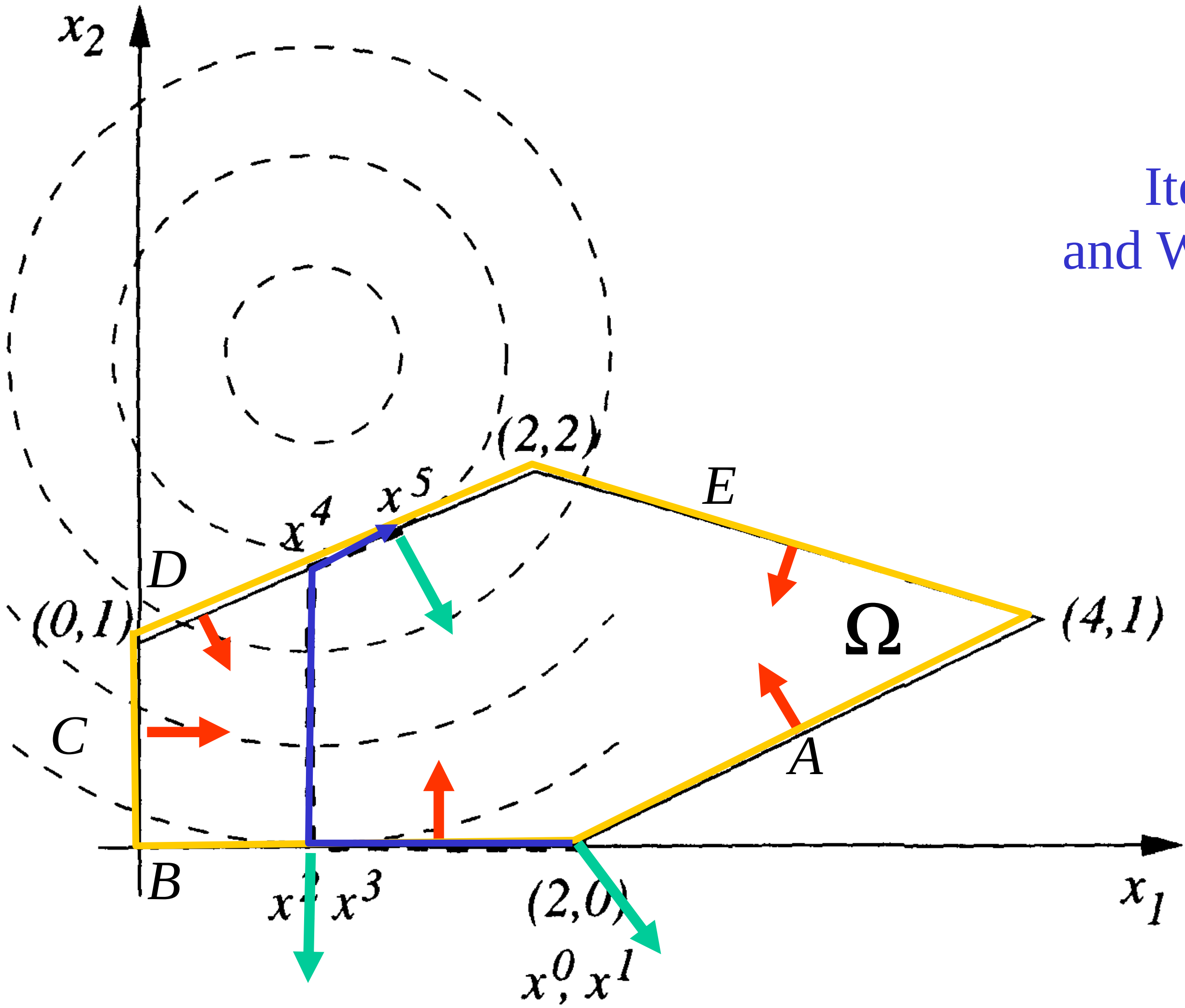
$$x_1 - 2x_2 + 2 \geq 0,$$

$$-x_1 - 2x_2 + 6 \geq 0,$$

$$-x_1 + 2x_2 + 2 \geq 0,$$

$$x_1 \geq 0,$$

$$x_2 \geq 0.$$



Iterations
and Working
Sets:
1: *A*,*B*
2: *B*
3: *B*
4: Ø
5: *D*
6: *D*

THE GRADIENT PROJECTION METHODS

The traditional gradient projection method admits non-linear objective functions as long as the constraints are linear:

$$\min f(x),$$
$$a_i' x = b_i,\ i \in \mathcal{E},$$
$$a_i' x \geq b_i,\ i \in \mathcal{I}.$$

We are in a point $x_k \in \Omega$ with a corresponding set of active constraints $\mathcal{A}_k$ and the (full rank) matrix of gradients A_k .

A (local) feasible domain is then

$$\Omega_k = \left\{x\ ;\ A_k x = b_k\right\} = \{x_k + \mathcal{N}(A_k)\}.$$

The gradient in x_k is

$$g_k = \nabla f(x_k)',$$

but in general $x_k - \alpha g_k$ will not be in Ω_k for any $\alpha \neq 0$.

We therefore project the gradient onto $\mathcal{N}(A_k)$ and consider the 1-D problem

$$\min_{\alpha} f(x_k - \alpha P_{\mathcal{N}(A_k)} g_k),$$

$$x_k - \alpha P_{\mathcal{N}(A_k)} g_k \in \Omega.$$

$$P_{\mathcal{N}(A_k)} = I - A_k'(A_k A_k')^{-1} A_k$$

We find the operator by solving the equality constrained QP-problem

$$\min_x \|g - x\|^2,$$

$$Ax = 0.$$

THE NON-LINEAR PROJECTION METHOD

$$\min q(x),$$
$$l \leq x \leq u.$$

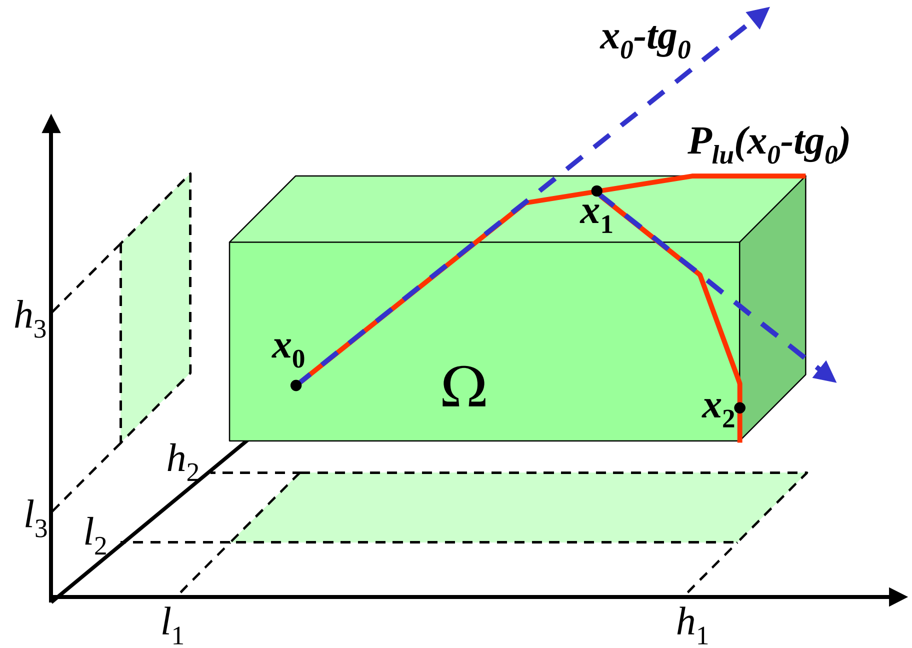
Consider the following (and obvious!) non-linear projection operator

$$P_{lu}\left(x\right)_i = \begin{cases} l_i, & x_i \leq l_i \\ x_i, & l_i < x_i < u_i \\ u_i, & u_i \leq x_i \end{cases}$$

We start at x_0 and compute the continuous broken line path

$$x(t) = P_{lu}(x_0 - t\nabla q(x_0)).$$

$$x(t) \in \Omega$$



- Let x_c be the first local minimum along the path.
- From this point the simplest would be to just compute a new gradient and repeat the operation.
- It is also possible to possible to do an approximate Active Set iteration using the already active bounds as the active set, as discussed in N&W p. 480.