Contents

Chapter 1

Quadratic Programming

1.1 Introduction

A *quadratic program (QP)* is an optimization problem where the objective function is quadratic and the constraints are linear. Problems of this type are important in their own right, and they also arise as subproblems in methods for general constrained optimization, such as sequential quadratic programming (Nocedal and Wright, 1999).

A general quadratic programming problem can be stated as:

$$\min_{\mathbf{x}} f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} \tag{1.1}$$

subject to

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1.2}$$

$$\mathbf{D}\mathbf{x} \leq \mathbf{g} \tag{1.3}$$

where:

$\mathbf{x}$	is a $n \times 1$ vector: $\mathbf{x} = [x_1 \;\; x_2 \;\; \ldots x_n]^T$
$\mathbf{Q}$	is a symmetric $n \times n$ matrix
$\mathbf{c}$	is an $n \times 1$ vector
$\mathbf{A}$	is a $m \times n$ matrix
$\mathbf{b}$	is an $m \times 1$ vector
$\mathbf{D}$	is a $p \times n$ matrix
$\mathbf{g}$	is an $p \times 1$ vector

The problem has n decision variables, m equality constraints and p inequality constraints.

If the matrix $\mathbf{Q}$ is positive semidefinite, the problem is a *convex quadratic program*.

Example 1.1 *(Nocedal and Wright, 1999) Consider the following QP problem:*

$$\min_{\mathbf{x}} f(\mathbf{x}) = 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3 \quad (1.4)$$

subject to

$$\begin{aligned} x_1 + x_3 &= 3 \quad &(1.5)\\ x_2 + x_3 &= 0 \quad &(1.6) \end{aligned}$$

We shall write the problem in the general form (1.1 -1.2).

In general, a quadratic form $\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x}$*, where* $\mathbf{Q}$ *is a symmetric* 3×3 *matrix and* $\mathbf{x}$ *is a* 3×1 *vector, can be written as:*

$$\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} = \frac{1}{2}[x_1 \;\; x_2 \;\; x_3] \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \frac{1}{2}(q_{11}x_1^2 + q_{22}x_2^2 + q_{33}x_3^2 + 2q_{12}x_1x_2 + 2q_{13}x_1x_3 + 2q_{23}x_2x_3)$$

$$= \frac{1}{2}q_{11}x_1^2 + \frac{1}{2}q_{22}x_2^2 + \frac{1}{2}q_{33}x_3^2 + q_{12}x_1x_2 + q_{13}x_1x_3 + q_{23}x_2x_3 \quad (1.7)$$

By simple identification of the coefficients, the objective function is written as a quadratic form:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \begin{bmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{x} + [-8 \;\; -3 \;\; -3]\mathbf{x} \quad (1.8)$$

The equality constraints, in the form $\mathbf{A}\mathbf{x} = \mathbf{b}$*, result as:*

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad (1.9)$$

Thus, the matrices $\mathbf{Q}$, $\mathbf{A}$ *and the vectors* $\mathbf{c}$ *and* $\mathbf{b}$ *are:*

$$\mathbf{Q} = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -8 \\ -3 \\ -3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \tag{1.10}$$

1.2 Equality-constrained quadratic programming

We consider first the case when the problem has only equality constraints. The QP-problem is stated as:

$$\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} \tag{1.11}$$

subject to

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1.12}$$

where $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{b}$ an $m \times 1$ vector and $\mathbf{x}$ an $n \times 1$ vector, with $m \leq n$.

If the number of constraints (m) is equal to the number of variables (n) the problem is uniquely determined and the only solution is the one of the linear system (1.12). In case the number of constraints is greater than the number of variables, $m \geq n$, the problem may have no solution.

The method of Lagrange multipliers will be applied. Let λ be the vector of m Lagrange multipliers, $\lambda = [\lambda_1 \quad \lambda_2 \quad \ldots \lambda_m]^T$. The Lagrangian of this problem is then:

$$L(\mathbf{x}, \lambda) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} + \lambda^T(\mathbf{A}\mathbf{x} - \mathbf{b}) \tag{1.13}$$

The necessary condition for a vector $\mathbf{x}$ to be a solution of the QP problem is that there is a vector λ such that the derivatives of the Lagrangian with respect to $\mathbf{x}$ and λ are zero:

$$\frac{\partial L}{\partial \mathbf{x}} = \mathbf{Q}\mathbf{x} + \mathbf{c} + \mathbf{A}^T\lambda = 0 \tag{1.14}$$

and

$$\frac{\partial L}{\partial \lambda} = \mathbf{A}\mathbf{x} - \mathbf{b} = 0 \tag{1.15}$$

The equations (1.14), (1.15) can be written as a system of linear equations:

$$\begin{aligned} \mathbf{Q}\mathbf{x} + \mathbf{A}^T\lambda &= -\mathbf{c} \\ \mathbf{A}\mathbf{x} - \mathbf{0}\cdot\lambda &= \mathbf{b} \end{aligned} \tag{1.16}$$

or:

$$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \mathbf{b} \end{bmatrix} \tag{1.17}$$

We shall write again the equation (1.17) emphasizing the size of the matrices involved:

$$\begin{bmatrix} \mathbf{Q}_{n\times n} & \mathbf{A}^T_{n\times m} \\ \mathbf{A}_{m\times n} & \mathbf{0}_{m\times m} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_{n\times 1} \\ \lambda_{m\times 1} \end{bmatrix} = \begin{bmatrix} -\mathbf{c}_{n\times 1} \\ \mathbf{b}_{m\times 1} \end{bmatrix} \tag{1.18}$$

The problem is reduced to the solution of a linear system of $n+m$ equations with $n+m$ unknowns. The sufficient conditions for minimum can be verified using, for example, the procedure described in section **??**.

Example 1.2 *Solve the following QP problem:*

$$\min_{(x_1,x_2)} f(x_1,x_2) = x_1^2 + x_2^2 \tag{1.19}$$

subject to

$$x_1 + x_2 = 5 \tag{1.20}$$

The objective function and the constraint can be written in the general matrix form as:

$$f(x_1,x_2) = f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \mathbf{x} + [0 \;\; 0]^T\mathbf{x} \tag{1.21}$$

$$[1 \;\; 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5 \tag{1.22}$$

The necessary condition for optimum is obtained from the linear system (1.17) given by:

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \tag{1.23}$$

and has the solution:

$$x_1 = 2.5, \quad x_2 = 2.5, \quad \lambda = -5 \tag{1.24}$$

*The sufficient condition for minimum is given by the sign of the determinant (**??**) which, in this example, is:*

$$(-1)^1 \begin{vmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 4 > 0 \tag{1.25}$$

The point is a unique minimizer of the objective function subject to the given constraint, because the determinant is positive.

Figure 1.1 shows the contour plot of the objective function and the constraint. The solution is the tangent point of one level curve and the constraint line.

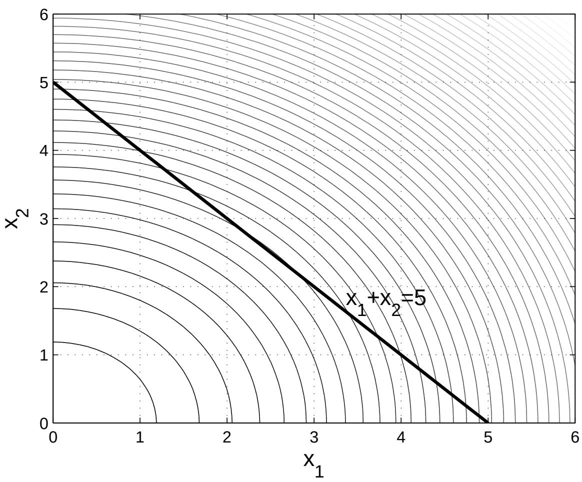

Figure 1.1: The contour plot of $f(x_1, x_2)$ and the constraint

1.3 Active set method for quadratic programming problems

1.3.1 Problem formulation

The general QP problem is stated now in a general form which emphasizes the m equality constraints and p inequality constraints:

$$\min_{\mathbf{x}} \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} \tag{1.26}$$

subject to

$$\begin{aligned} \mathbf{a}_i^T\mathbf{x} &= b_i, \quad i = \overline{1,m} & (1.27)\\ \mathbf{a}_j^T\mathbf{x} &\leq b_j, \quad j = \overline{m+1, m+p} & (1.28) \end{aligned}$$

where:

$\mathbf{x}$	is the $n \times 1$ vector of unknowns
$\mathbf{Q}$	is a symmetric $n \times n$ positive semidefinite matrix
$\mathbf{c}$	is an $n \times 1$ vector
b_i	are scalars, $i = \overline{1, m+p}$
$\mathbf{a}_i$	are $n \times 1$ vectors, $i = \overline{1, m+p}$

The vectors $\mathbf{a}_i^T$ are the rows of the matrices $\mathbf{A}$ and $\mathbf{D}$ from (1.2) and (1.3).

We define the *active set*, $\mathcal{A}(\mathbf{x}^*)$, at an optimal point $\mathbf{x}^*$ as the indices of the constraints at which equality holds, that is (Nocedal and Wright, 1999):

$$\mathcal{A}(\mathbf{x}^*) = \{i \in \{1, 2, \ldots, m+p\} : \mathbf{a}_i^T\mathbf{x}^* = b_i\} \tag{1.29}$$

The Lagrangian of this problem is:

$$L(\mathbf{x}, \lambda) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} + \sum_{i=1}^{m}\lambda_i(\mathbf{a}_i^T\mathbf{x} - b_i) + \sum_{j=m+1}^{m+p}\lambda_j(\mathbf{a}_j^T\mathbf{x} - b_j) \tag{1.30}$$

where λ is a vector of $m+p$ Lagrange multipliers.

We can write now the KKT conditions for this problem:

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{x}} &= 0 && (1.31)\\
\mathbf{a}_i^T\mathbf{x} &= b_i,\ \ i = \overline{1,m} && (1.32)\\
\mathbf{a}_j^T\mathbf{x} &\leq b_j,\ \ j = \overline{m+1, m+p} && (1.33)\\
\lambda_j(\mathbf{a}_j^T\mathbf{x} - b_j) &= 0,\ \ j = \overline{m+1, m+p} && (1.34)\\
\lambda_j &\geq 0,\ \ j = \overline{m+1, m+p} && (1.35)\\
\lambda_i &\ \ ,\ \ i = \overline{1,m} \text{ unrestricted in sign} && (1.36)
\end{aligned}$$

Example 1.3 *Consider the following QP problem:*

$$\min_{(x_1,x_2)} f(x_1, x_2) = x_1^2 + 2x_2^2 - x_1 - 2x_2 \tag{1.37}$$

subject to

$$\begin{aligned}
x_1 + x_2 &= 3\\
x_1 &\geq 1\\
x_2 &\leq 3\\
-x_1 + 3x_2 &\geq -1\\
x_1 + x_2 &\leq 5
\end{aligned} \tag{1.38}$$

We shall illustrate the feasible region and write the problem in the general form (1.26), (1.27), (1.28).

The inequality constraints form the convex gray region illustrated in Figure 1.2. The minimizer of the objective function subject to the constraints must be located in the interior of the feasible region, on the straight line $x_1 + x_2 = 3$ (the equality constraint).

If we denote $\mathbf{x} = [x_1 \quad x_2]^T$, the quadratic objective function is written in the matrix form as:

$$\begin{aligned}
f(\mathbf{x}) &= \frac{1}{2}\mathbf{x}^T \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \mathbf{x} + [-1 \quad -2]\mathbf{x}\\
&= \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} && (1.39)
\end{aligned}$$

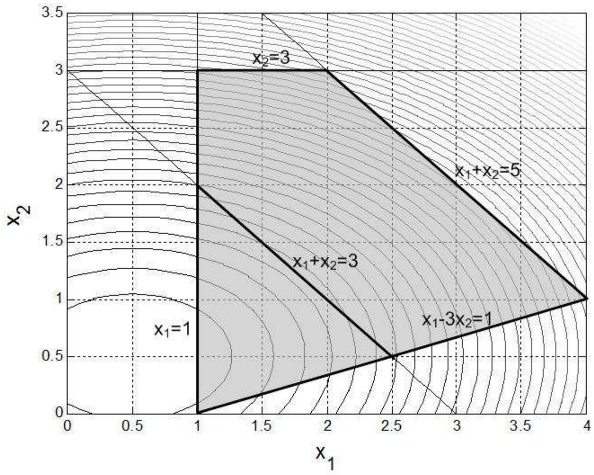

Figure 1.2: The feasible region and the contour plot of the objective function

The number of equality constraints for this problem is $m = 1$. *It can be written as:*

$$x_1 + x_2 = [1 \;\; 1]\,\mathbf{x} = 3 \tag{1.40}$$

thus:

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } b_1 = 3 \tag{1.41}$$

Some of the inequality constraints will be multiplied by (-1) *so all the left hand sides are less or equal than the right hand side:*

$$\begin{aligned} -x_1 &\leq -1 \\ x_2 &\leq 3 \\ x_1 - 3x_2 &\leq 1 \\ x_1 + x_2 &\leq 5 \end{aligned} \tag{1.42}$$

The vectors $\mathbf{a}_j,\ j = 2,\ 3,\ 4,\ 5$ *are:*

$$\mathbf{a}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix},\ \mathbf{a}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ \mathbf{a}_4 = \begin{bmatrix} 1 \\ -3 \end{bmatrix},\ \mathbf{a}_5 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{1.43}$$

and the scalars from the right hand side:

$$b_2 = -1, \;\; b_3 = 3, \;\; b_4 = 1, \;\; b_5 = 5 \tag{1.44}$$

We shall not give a solution for this problem, but it is clear from the Figure 1.2 that the optimum is located in the interior of the feasible region, at some point where the equality-constraint line is tangent to a level curve of the objective function. Since at this point, all the inequality constraints are inactive (none of them is satisfied as an equality), the optimal active set is $\mathcal{A}(\mathbf{x}^*) = \{1\}$*, which is the index of the only constraint fulfilled as equality.*

1.3.2 The active set method

The active set method is one of the most popular approaches for solving small and medium scale QP problems. The idea behind the method may be summarized as follows (Nocedal and Wright, 1999; Bhatti, 2000):

- Start with a guess of the optimal active set $\mathcal{A}$ and calculate a feasible initial iterate $\mathbf{x}_0$.
- Use the gradient and Lagrange multiplier information to remove one index from the current active set and to add a new one. The method ensures the feasibility of the next iterate $\mathbf{x}_{k+1}$ calculated from:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \tag{1.45}$$

  where α_k is the step length and $\mathbf{d}_k$ the direction of moving, obtained by solving a QP subproblem. This subproblem will have a subset of constraints imposed as equalities and referred as *the working set*, W_k, consisting of all m equality constraints and some of the active inequalities. Some iterates may be located on the boundary or in the interior of the feasible region.
- New iterates are calculated and the working set is modified until the optimality conditions are satisfied, or all Lagrange multipliers are positive as required by the KKT conditions.

Let $\mathbf{x}_k$ be the current iterate. At this point, some of the inequality constraints may be active (or satisfied as equalities). Together with the equality constraints they form the working set W_k:

$$W_k = \{1, \ldots, m\} \cup \{i : \mathbf{a}_i^T \mathbf{x}_k = b_i,\ i = m+1, \ldots, m+p\} \tag{1.46}$$

For the current point, we check whether $\mathbf{x}_k$ minimizes the quadratic objective function in the subspace defined by the working set, i.e. the Lagrange multipliers corresponding to the inequality constraints are positive. This is a direct consequence of the KKT conditions.

If the optimality conditions are not satisfied, we compute a direction, $\mathbf{d}_k$, to move to the next point $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$ such that the new iterate is feasible in W_k and the objective function is minimized at $\mathbf{x}_k + \mathbf{d}_k$. Since $\mathbf{x}_k$ is known at the current stage, it will be regarded as a constant vector and the unknown vector is only $\mathbf{d}_k$. The problem is stated as:

$$\min_{\mathbf{d}_k} f(\mathbf{d}_k) = \frac{1}{2}(\mathbf{x}_k + \mathbf{d}_k)^T \mathbf{Q} (\mathbf{x}_k + \mathbf{d}_k) + \mathbf{c}^T (\mathbf{x}_k + \mathbf{d}_k) \tag{1.47}$$

subject to:

$$\mathbf{a}_i^T (\mathbf{x}_k + \mathbf{d}_k) = b_i, \quad i \in W_k \tag{1.48}$$

Expanding the new objective function we have:

$$f(\mathbf{d}_k) = \frac{1}{2}\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \frac{1}{2}\mathbf{d}_k^T \mathbf{Q} \mathbf{d}_k + \mathbf{x}_k^T \mathbf{Q} \mathbf{d}_k + \mathbf{c}^T \mathbf{x}_k + \mathbf{c}^T \mathbf{d}_k \tag{1.49}$$

The term $\frac{1}{2}\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + \mathbf{c}^T \mathbf{x}_k$ is constant for a given $\mathbf{x}_k$, thus it can be removed from the objective function without changing the solution.

We denote:

$$\mathbf{g}_k = \mathbf{Q} \mathbf{x}_k + \mathbf{c} \tag{1.50}$$

and the function to be minimized becomes:

$$f(\mathbf{d}_k) = \frac{1}{2}\mathbf{d}_k^T \mathbf{Q} \mathbf{d}_k + (\mathbf{x}_k^T \mathbf{Q} + \mathbf{c}^T)\mathbf{d}_k = \frac{1}{2}\mathbf{d}_k^T \mathbf{Q} \mathbf{d}_k + \mathbf{g}_k^T \mathbf{d}_k \tag{1.51}$$

Note that $\mathbf{Q}$ is symmetric, thus $\mathbf{Q} = \mathbf{Q}^T$.

Because $\mathbf{x}_k$ is a feasible point within the working set W_k, the equality constraint:

$$\mathbf{a}_i^T \mathbf{x}_k = b_i, \quad i \in W_k \tag{1.52}$$

is satisfied. From (1.52) and (1.48) we obtain the equality constraint of the new QP subproblem. It will be formulated as:

$$\min_{\mathbf{d_k}} \frac{1}{2}\mathbf{d}_k^T\mathbf{Q}\mathbf{d}_k + \mathbf{g}_k^T\mathbf{d}_k \tag{1.53}$$

subject to:

$$\mathbf{a}_i^T\mathbf{d}_k = 0, \quad i \in W_k \tag{1.54}$$

We may proceed in a manner similar to the one applied for equality-constrained QP problems. If we denote by $\mathbf{A}$ the matrix having the rows $\mathbf{a}_i^T$, for all indices i in the working set:

$$\mathbf{A} = \left[\mathbf{a}_i^T\right], \quad i \in W_k \tag{1.55}$$

and $\mathbf{b}$ the vector of all $b_i,\ i \in W_k$, the Lagrangian is:

$$L(\mathbf{d}_k, \lambda) = \frac{1}{2}\mathbf{d}_k^T\mathbf{Q}\mathbf{d}_k + \mathbf{g}_k^T\mathbf{d}_k + \lambda^T\mathbf{A}\mathbf{d}_k \tag{1.56}$$

The first-order optimality condition result as:

$$\frac{\partial L}{\partial \mathbf{d}_k} = \mathbf{Q}\mathbf{d}_k + \mathbf{g}_k + \mathbf{A}^T\lambda = 0 \tag{1.57}$$

$$\mathbf{A}\mathbf{d}_k = 0 \tag{1.58}$$

and can be written in a matrix form:

$$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{d}_k \\ \lambda \end{bmatrix} = \begin{bmatrix} -\mathbf{g}_k \\ 0 \end{bmatrix} \tag{1.59}$$

The direction $\mathbf{d}_k$ and the multipliers are the solution of the linear system (1.59).

- If $\mathbf{d}_k$ is nonzero we shall move in this direction. Note that for all indices i in the working set W_k, the term $\mathbf{a}_i^T\mathbf{x}_k$ does not change as we move along $\mathbf{d}_k$ because:

$$\mathbf{a}_i^T(\mathbf{x}_k + \mathbf{d}_k) = \mathbf{a}_i^T\mathbf{x}_k = b_i \text{ and from (1.54): } \mathbf{a}_i^T\mathbf{d}_k = 0 \tag{1.60}$$

Since the constraints in W_k were satisfied at $\mathbf{x}_k$, they are satisfied at $\mathbf{x}_k + \alpha_k \mathbf{d}_k$, for any value of α.

If the point $\mathbf{x}_k + \mathbf{d}_k$ is feasible for all constraints (including the ones that are not in W_k), we compute the next point: $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$. Otherwise, we set $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$, where the step length is chosen as the largest value in $[0, 1)$ to satisfy all the constraints.

Because the constraints in W_k are satisfied for any α_k, a relation to compute the step length will be determined such that $\mathbf{x}_k + \alpha_k \mathbf{d}_k$ satisfies the constraints $i \notin W_k$.

– If $\mathbf{a}_i^T \mathbf{d}_k \leq 0$ for some constraints $i \notin W_k$, the feasibility is satisfied for any value of α_k because:

$$\mathbf{a}_i^T (\mathbf{x}_k + \alpha_k \mathbf{d}_k) = \mathbf{a}_i^T \mathbf{x}_k + \alpha_k \mathbf{a}_i^T \mathbf{d}_k \leq \mathbf{a}_i^T \mathbf{x}_k \leq b_i \tag{1.61}$$

– If $\mathbf{a}_i^T \mathbf{d}_k > 0$ for some constraints $i \notin W_k$, the value of α_k is calculated from:

$$\mathbf{a}_i^T (\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq b_i \tag{1.62}$$

or:

$$\alpha_k \leq \frac{b_i - \mathbf{a}_i^T \mathbf{x}_k}{\mathbf{a}_i^T \mathbf{d}_k} \tag{1.63}$$

Because we want α_k to be as large as possible in the interval $[0; 1]$, it will be calculated from (Nocedal and Wright, 1999; Bhatti, 2000):

$$\alpha_k = \min_{i \notin W_k,\, \mathbf{a}_i^T \mathbf{d}_k > 0} \left(1, \frac{b_i - \mathbf{a}_i^T \mathbf{x}_k}{\mathbf{a}_i^T \mathbf{d}_k} \right) \tag{1.64}$$

If $\alpha_k < 1$, the movement along $\mathbf{d}_k$ was blocked by some constraint that does not belong to W_k. This is a *blocking constraint* and will be added to the new working set.

If $\alpha_k = 1$ no new constraints are active for $\mathbf{x}_k + \alpha_k \mathbf{d}_k$ and there are no blocking constraints at this stage.

- If $\mathbf{d}_k$ is zero we have to check for optimality, i.e. all the Lagrange multipliers for the inequality constraints must be positive. If this is not the case, and some of them are negative, the constraint corresponding to the multiplier having the smallest negative value will be removed from the working set.

The iterations are continued until we reach a point that minimizes the objective over the current working set. In this case the directions $\mathbf{d}_k$ will be zero and the Lagrange multipliers are positive.

An algorithm for the active set method is described in Algorithm 1.

Example 1.4 *Solve the following QP problem using the active set method:*

$$\min_{x_1,x_2} f(x_1,x_2) = x_1^2 + x_2^2 - 4x_1 - 4x_2 \tag{1.69}$$

subject to

$$\begin{aligned} (1):& \quad x_1 + x_2 &\leq& \; 2\\ (2):& \quad x_1 - 2x_2 &\leq& \; 2\\ (3):& \quad -x_1 - x_2 &\leq& \; 1\\ (4):& \quad -2x_1 + x_2 &\leq& \; 2 \end{aligned} \tag{1.70}$$

The problem is written in the quadratic matrix form:

$$\begin{aligned} \min_{\mathbf{x}} f(\mathbf{x}) &= \frac{1}{2}\mathbf{x}^T \begin{bmatrix} 2 & 0\\ 0 & 2 \end{bmatrix} \mathbf{x} + [-4 \quad -4]\mathbf{x}\\ &= \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} \end{aligned} \tag{1.71}$$

subject to

$$\begin{aligned} (1):& \quad \begin{bmatrix} 1 & 1 \end{bmatrix}\mathbf{x} \leq 2\\ (2):& \quad \begin{bmatrix} 1 & -2 \end{bmatrix}\mathbf{x} \leq 2\\ (3):& \quad \begin{bmatrix} -1 & -1 \end{bmatrix}\mathbf{x} \leq 1\\ (4):& \quad \begin{bmatrix} -2 & 1 \end{bmatrix}\mathbf{x} \leq 2 \end{aligned} \tag{1.72}$$

Algorithm 1 Active set method

Define the quadratic objective function: matrix $\mathbf{Q}$ and vector $\mathbf{c}$
Define the constraints: vectors $\mathbf{a}_i^T$ and scalars b_i, $i = \overline{1,p}$
Select an initial feasible point $\mathbf{x}_0$
Find the initial working set W_0
Compute the gradient of the objective function at the current point: $\mathbf{g}_0 = \mathbf{Q}\mathbf{x}_0 + \mathbf{c}$
Compute the matrix $\mathbf{A}$ having the rows $\mathbf{a}_i$, $i \in W_k$
Solve the linear system

$$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \lambda \end{bmatrix} = \begin{bmatrix} -\mathbf{g}_0 \\ \mathbf{0} \end{bmatrix} \tag{1.65}$$

Set $k = 0$
while not all $\lambda_i \geq 0,\ i \in W_k \cap \{1, 2, ..., m\}$ or $\mathbf{d}_k \neq 0$ **do**
  if $\mathbf{d}_k = 0$ **then**
    Check optimality:
    if $\lambda_i \geq 0$, $i \in W_k \cap \{1, 2, ..., m\}$ **then**
      Stop and return the current point $\mathbf{x}_k$
    else
      Find the most negative λ_j
      Remove constraint j from the working set W_k
      Keep the same point for the next step: $\mathbf{x}_{k+1} = \mathbf{x}_k$
    end if
  else
    Compute the step length α_k from:

$$\alpha_k = \min_{i \notin W_k,\, \mathbf{a}_i^T \mathbf{d}_k > 0} \left(1, \frac{b_i - \mathbf{a}_i^T \mathbf{x}_k}{\mathbf{a}_i^T \mathbf{d}_k}\right) \tag{1.66}$$

    Compute a new point: $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$
    If $\alpha_k < 1$, find the blocking constraint i and add it to the working set W_k
  end if
  Set $k \leftarrow k + 1$
  Compute the gradient of the objective function at the current point:

$$\mathbf{g}_k = \mathbf{Q}\mathbf{x}_k + \mathbf{c} \tag{1.67}$$

  Compute the matrix $\mathbf{A}$ having the rows $\mathbf{a}_i$, $i \in W_k$
  Solve the linear system

$$\begin{bmatrix} \mathbf{Q} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{d}_k \\ \lambda \end{bmatrix} = \begin{bmatrix} -\mathbf{g}_k \\ \mathbf{0} \end{bmatrix} \tag{1.68}$$

end while

The vectors $\mathbf{a}_i^T$, $i = \overline{1,4}$ *are:*

$$\mathbf{a}_1^T = \begin{bmatrix} 1 & 1 \end{bmatrix} \tag{1.73}$$

$$\mathbf{a}_2^T = \begin{bmatrix} 1 & -2 \end{bmatrix} \tag{1.74}$$

$$\mathbf{a}_3^T = \begin{bmatrix} -1 & -1 \end{bmatrix} \tag{1.75}$$

$$\mathbf{a}_4^T = \begin{bmatrix} -2 & 1 \end{bmatrix} \tag{1.76}$$

- *We select as the starting point* $\mathbf{x}_0 = [0 \quad -1]^T$. *The objective function, the constraints and the iterates are illustrated in Figure 1.3. The feasible region is the interior and the sides of the polygon bordered by the constraints numbered as in the relations (1.70).*

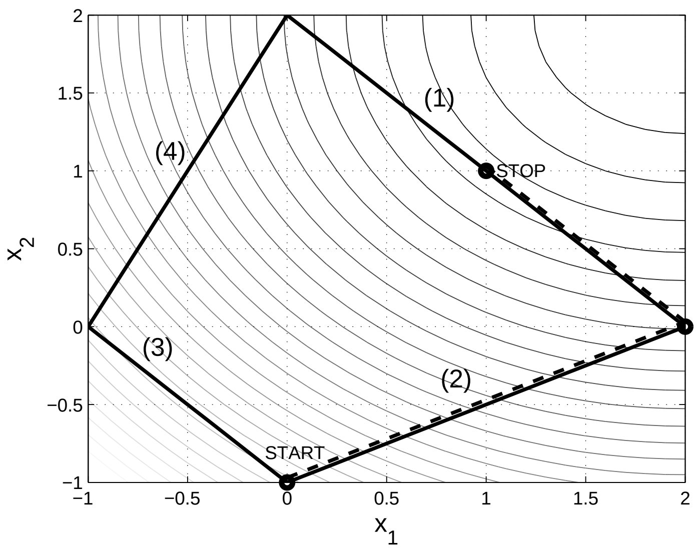

Figure 1.3: Active set method with $\mathbf{x}_0 = [0 \quad -1]^T$ and $W_0 = \{2,\ 3\}$

For the initial point selected as the vertex (0, -1), (or as the intersection of the lines $x_1 = 2x_2 = 2$ *and* $-x_1 - x_2 = 1$*), the constraints (2) and (3) are active. The initial working set is then* $W_0 = \{2,\ 3\}$. *We may choose* $W_0 = \{2\}$, *or* $\{3\}$, *or an empty set, and the algorithm will return the same result, but the iterates will follow another trajectory.*

The rows of matrix $\mathbf{A}$ *are* $\mathbf{a}_i^T,\ i = \{2, 3\}$*:*

$$\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -1 & -1 \end{bmatrix} \tag{1.77}$$

The gradient $\mathbf{g}_0$ *is obtained from* $\mathbf{Q}$*,* $\mathbf{c}$ *and* $\mathbf{x}_0$ *and the solution of the linear system (1.65) is:*

$$\mathbf{d}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} -0.6667 \\ -4.6667 \end{bmatrix} \tag{1.78}$$

Step 1. *The vector* $\mathbf{d}_0$ *is zero, thus we check optimality: all* λ*-multipliers are negative. We remove the constraint (3) from the working set since it has the most negative multiplier and obtain* $W_1 = \{2\}$*. The matrix* $\mathbf{A}$ *is now:* $\mathbf{A} = [1 \quad -2]$ *and the linear system (1.68) has the solution:*

$$\mathbf{d}_1 = \begin{bmatrix} 2.8 \\ 1.4 \end{bmatrix}, \quad \lambda = -1.6 \tag{1.79}$$

Step 2. *The vector* $\mathbf{d}_1$ *is nonzero. The step length formula (1.66) yields:* $\alpha = 0.7143$*. The only positive product* $\mathbf{a}_i^T\mathbf{d}_1$ *where* $i \notin W_k$ *is obtained for* $i = 1$*. The first constraint will be added to the working set and* $W_2 = \{2,\ 1\}$*.*

The new iterate is:

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha\mathbf{d}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \tag{1.80}$$

and the solution of (1.68) result as:

$$\mathbf{d}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \lambda = \begin{bmatrix} -1.3333 \\ 1.3333 \end{bmatrix} \tag{1.81}$$

Step 3. *The vector* $\mathbf{d}_2$ *is zero, thus we check optimality: the first multiplier is negative and the constraint (2) will be removed from the working set which becomes:* $W_3 = \{1\}$*. The iterate* $\mathbf{x}_3$ *is the same as in the previous step, the matrix* $\mathbf{A} = [1 \ \ 1]$ *and the linear system (1.68) has the solution:*

$$\mathbf{d}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \lambda = 2 \tag{1.82}$$

Step 4. *The vector* $\mathbf{d}_1$ *is nonzero although the multiplier is positive. The step length formula (1.66) yields:* $\alpha = 1$. *The working set in the same as before* $W_4 = \{1\}$ *and the new iterate is:*

$$\mathbf{x}_4 = \mathbf{x}_3 + \mathbf{d}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{1.83}$$

The linear system (1.68) has the solution:

$$\mathbf{d}_4 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \lambda = 2 \tag{1.84}$$

Step 5. *The vector* $\mathbf{d}_4$ *is zero and the optimality condition is satisfied (*$\lambda = 2 > 0$*). The minimizer of the objective function subject to the given constraints is* $\mathbf{x}^* = [1 \quad 1]^T$.

The steps taken in this procedure are shown in Figure 1.3. The same result is obtained if the initial point is changed to another feasible location, or if we choose another initial working set, as shown in the following cases:

- *We select as the starting point* $\mathbf{x}_0 = [-0.2 \quad -0.8]^T$ *and the initial working set* $W_0 = \{3\}$. *The objective function, the constraints and the iterates are illustrated in Figure 1.4.*
- *If the initial point is* $\mathbf{x}_0 = [0 \quad -1]^T$ *and the initial working set* $W_0 = \{3\}$, *the result is the same as in previous cases, as shown in Figure 1.5.*
- *Another test was performed for an initial point* $\mathbf{x}_0 = [0 \quad -1]^T$ *and an empty working set* $W_0 = \{\emptyset\}$. *The trajectory of the iterates is different from previous cases, but the final point is the same as shown in Figure 1.6.*

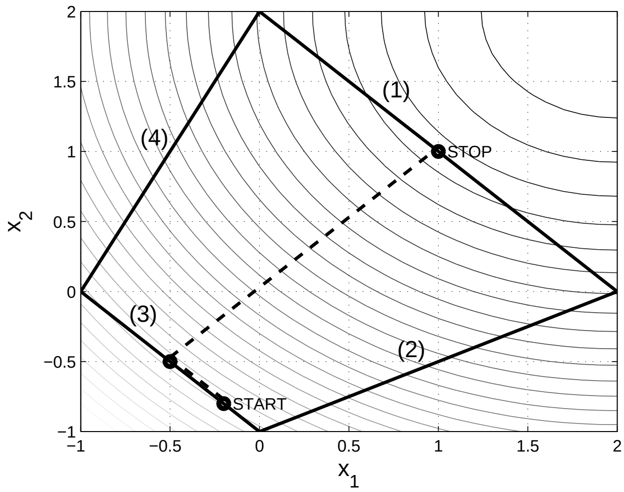

Figure 1.4: Active set method with $\mathbf{x}_0 = [-0.2 \quad -0.8]^T$ and $W_0 = \{3\}$

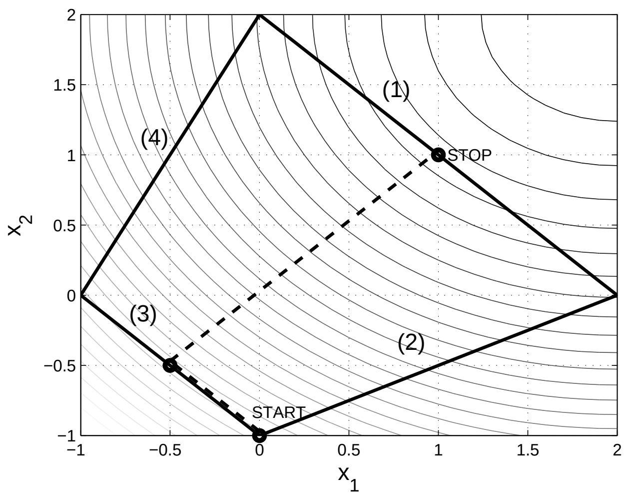

Figure 1.5: Active set method with $\mathbf{x}_0 = [0 \quad -1]^T$ and $W_0 = \{3\}$

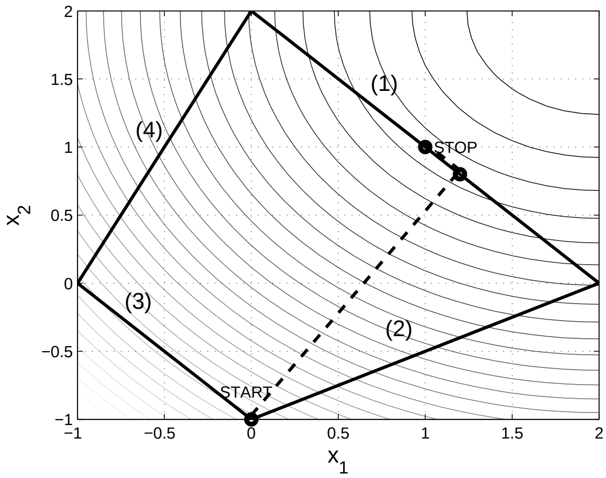

Figure 1.6: Active set method with $\mathbf{x}_0 = [0 \quad -1]^T$ and $W_0 = \{\emptyset\}$

Bibliography

Bhatti, M. A. (2000). *Practical Optimization Methods: With Mathematica Applications*. Springer.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. Springer series in operation research. Springer-Verlag.