STOR 614 Handout

The active set method for quadratic programs

Equality-constrained quadratic programs

First, consider a quadratic program with only equality constraints:

$$\begin{aligned} &\min\ q(x) = \frac{1}{2}x^T G x + c^T x \\ &\text{s.t. } Ax = b \end{aligned} \tag{1}$$

with $G \in \mathbb{R}^{n\times n}$ symmetric, and $A \in \mathbb{R}^{m\times n}$ written as

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}.$$

The KKT conditions for (1) can be written as

$$\begin{aligned} &Gx^* + c - \sum_{i=1}^{m} \lambda_i^* a_i = 0, \\ &Ax^* = b \end{aligned} \tag{2}$$

or equivalently

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

Theorem 1. *Let Z be a matrix whose columns form a basis for the null space of A. If A is of full row rank, and $Z^T G Z$ is positive definite, then*

1. *the matrix $\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix}$ is nonsingular, and there is a unique pair (x^*, λ^*) satisfying* (2)*;*

2. *x^* is the unique global solution for* (1)*.*

Proof. (1) It suffices to show the only vector (w, v) that satisfies

$$\begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = 0. \tag{3}$$

is $(w, v) = 0$. From (3) we have $Aw = 0$ and

$$\begin{bmatrix} w^T & v^T \end{bmatrix} \begin{bmatrix} G & -A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = w^T G w - w^T A^T v + v^T A w = w^T G w = 0.$$

By the definition of Z, $w = Zu$ for some u. It follows that

$$0 = w^T G w = u^T Z^T G Z u,$$

which implies $u = 0$ by the assumption $Z^T G Z$ is positive definite. Therefore, $w = Au = 0$. From (3) we also have $Gw - A^T v = 0$, which implies $A^T v = 0$. Since A is of full row rank, A^T is of full column rank. This shows that $v = 0$.

Proof of (2). Since the KKT conditions (2) are necessary to hold for any local solutions to (1), the only possible local solution to (1) is x^*. To show that x^* is actually a global solution, we can use a direct check. Let x satisfy $Ax = b$ and write $p = x - x^*$. Note that $Ap = 0$, which implies $p = Zu$ for some u. Then we have

$$\begin{aligned}
& q(x) - q(x^*) \\
=& \frac{1}{2}(x^* + p)^T G (x^* + p) + c^T (x^* + p) - \frac{1}{2} x^* G x^* - c^T x^* \\
=& \frac{1}{2} p^T G p + p^T (G x^* + c) \\
=& \frac{1}{2} p^T G p + p^T (A^T \lambda^*) \\
=& \frac{1}{2} p^T G p \\
=& \frac{1}{2} u^T Z^T G Z u > 0
\end{aligned}$$

whenever $u \neq 0$. □

Active set method for general quadratic programs

Now, consider a general quadratic program:

$$\begin{aligned}
\min\ & q(x) = \frac{1}{2} x^T G x + c^T x \\
\text{s.t.}\ & a_i^T x = b_i, \quad i \in \mathcal{E}; \\
& a_i^T x \geq b_i, \quad i \in \mathcal{I}.
\end{aligned} \tag{4}$$

In this section we make the following assumptions.

Assumption 1. *1. $G \in \mathbb{R}^{n \times n}$ is symmetric and positive definite.*

2. $\mathcal{E}$ and $\mathcal{I}$ are finite, disjoint index sets, and $\{a_i, i \in \mathcal{E}\}$ is linearly independent.

Recall the KKT conditions for (4) are

$$\begin{aligned}
& G x^* + c = \sum_{i \in \mathcal{E} \cup \mathcal{I}} a_i \lambda_i^* \\
& a_i^T x^* = b_i, \quad i \in \mathcal{E} \\
& b_i \leq a_i^T x^* \perp \lambda_i^* \geq 0, \quad i \in \mathcal{I}.
\end{aligned} \tag{5}$$

At the kth iteration of the algorithm, we have the current iterate x_k (a feasible solution), and a working index set W_k that satisfies the following two conditions:

1. $\mathcal{E} \subset W_k \subset A(x_k)$, where $A(x_k) = \{i \in \mathcal{E} \cup \mathcal{I} \mid a_i^T x_k = b_i\}$ is the active set of x_k.
2. $\{a_i,\ i \in W_k\}$ is linearly independent.

We then solve the following equality-constrained QP is solved, whose solution is denoted as p_k:

$$\begin{aligned} &\min_p \ q(x_k + p) = \frac{1}{2}p^T G p + (G x_k + c)^T p + (\frac{1}{2}x_k^T G x_k + c^T x_k) \\ &\text{s.t. } a_i^T p = 0, \quad i \in W_k. \end{aligned} \tag{6}$$

Depending on the solution p_k obtained, there are the following cases.

Case 1. If $p_k \neq 0$, then x_k is not the global solution of

$$\begin{aligned} &\min \ q(x) = \frac{1}{2}x^T G x + c^T x \\ &\text{s.t. } a_i^T x = b_i, \quad i \in W_k. \end{aligned} \tag{7}$$

Let

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

where α_k is the largest value in [0,1] such that x_{k+1} is feasible, and can be computed as

$$\alpha_k = \min\left(1, \min_{i \notin W_k,\ a_i^T p_k < 0} \frac{b_i - a_i^T x_k}{a_i^T p_k}\right).$$

We call α_k the *step length* of the kth iteration. Note that it is possible to have $\alpha_k = 0$ when $a_i^T p_k < 0$ for some $i \in A(x_k) \setminus W_k$. In addition, it follows from the fact $q(x_k + p_k) < q(x_k)$ and the convexity of q that $q(x_k + \alpha_k p_k) < q(x_k)$ whenever $\alpha_k > 0$.

Case 1.1. If $\alpha_k = 1$, then let $W_{k+1} = W_k$. (Note that we will have $p_{k+1} = 0$ in the next iteration because x_{k+1} as defined is the global solution of (7) with $W_{k+1} = W_k$.)

Case 1.2. If $\alpha_k < 1$, then there exists $i \notin W_k$ with $a_i^T p_k < 0$ such that

$$\alpha_k = \frac{b_i - a_i^T x_k}{a_i^T p_k}.$$

Let $W_{k+1} = \{i\} \cup W_k$. Note that $\mathcal{E} \subset W_{k+1} \subset A(x_{k+1})$, and $\{a_i, i \in W_{k+1}\}$ is linearly independent.

Case 2. If $p_k = 0$, then x_k is the global solution of (7). The KKT conditions of (7) imply that

$$Gx_k + c = \sum_{i \in W_k} \lambda_i a_i \tag{8}$$

for some $\lambda = (\lambda_i)_{i \in W_k}$.

Case 2.1. If $\lambda_i \geq 0$ for all $i \in W_k \cap I$, then x_k satisfies the KKT conditions for (4) and is therefore the global solution of (4). STOP.

Case 2.2. If $\lambda_j < 0$ for some $j \in W_k \cap I$, then let $x_{k+1} = x_k$ and $W_{k+1} = W_k \setminus \{j\}$.

Lemma 1. *If the kth iteration belongs to Case 2.2, then*

$$q(x_{k+1} + p_{k+1}) < q(x_{k+1}).$$

Proof. Consider the QP at the $k+1$th iteration:

$$\begin{aligned} \min_p \ & q(x_{k+1} + p) = \frac{1}{2} p^T G p + (Gx_k + c)^T p + (\frac{1}{2} x_k^T G x_k + c^T x_k) \\ \text{s.t.} \ & a_i^T p = 0, \quad i \in W_k \setminus \{j\}. \end{aligned} \tag{9}$$

Let p_{k+1} be the global solution to (9). It satisfies

$$Gx_k + c + Gp_{k+1} = \sum_{i \in W_k \setminus \{j\}} \tilde{\lambda}_i a_i.$$

Combining the above equality with (8) to get

$$Gp_{k+1} = \sum_{i \in W_k \setminus \{j\}} (\tilde{\lambda}_i - \lambda_i) a_i - \lambda_j a_j.$$

Recall that $\lambda_j < 0$. Since $\{a_i, i \in W_k\}$ is linearly independent, $p_{k+1} \neq 0$. Since $x_{k+1} + p_{k+1}$ is the unique global solution to (7) with W_{k+1} replacing W_k, the fact $p_{k+1} \neq 0$ implies that $q(x_{k+1} + p_{k+1}) < q(x_{k+1})$.

□

Theorem 2 (Finite termination of the active set method)**.** *Assume that Assumption 1 holds for the quadratic program* (4)*. Suppose additionally that the step lengths are strictly positive in all iterations of the active set method when applied to* (4)*. Then, the method finds the global solution to* (4) *in finite iterations.*

Proof. First, the method encounters Case 2 at least in every $n+1$ iterations. To see this, suppose that the kth iteration is in Case 1 (i.e., $p_k \neq 0$). If, for some $j = k, \cdots, k+(n-1)$, the jth iteration is in Case 1.1, then the $j+1$th iteration will be in Case 2. Otherwise, if for all $j = k, \cdots, k+(n-1)$, the

jth iteration is in Case 1.2, then by the end of the $k+(n-1)$th iteration, the working set W_{k+n} will contain n indices. Consequently we will have $p_{k+n} = 0$, and the $k+n$th iteration will belong to Case 2.

Next, if the kth and lth iterations are both in Case 2, then $W_k \neq W_l$. To see this, assume WLOG that $k < l$. Since the kth iteration is in Case 2.2, we will have $x_{k+1} = x_k$ and $p_{k+1} \neq 0$ (as shown in Lemma 1). Since $\alpha_{k+1} > 0$ by the assumption on step lengths, we have

$$q(x_{k+2}) = q(x_{k+1} + \alpha_{k+1} p_{k+1}) < q(x_{k+1}) = q(x_k).$$

This implies

$$q(x_l) \leq q(x_{k+2}) < q(x_k).$$

Since the kth and lth iterations are both in Case 2, x_k is the global solution of (7) and x_l is the global solution of (7) with W_k replaced by W_l. The fact $q(x_l) < q(x_k)$ thus implies $W_k \neq W_l$.

To summarize, the algorithm encounters Case 2 at least in every $n+1$ iterations, and a new working set is used in every Case 2 iteration. Since there are finitely many possible working sets, the algorithm must stop in finitely many iterations. □

For more details of implementation of the active set method, see [1].

Example 1. *Figure 1 shows the iteration points of an application of the active set method to the QP below:*

$$
\begin{aligned}
\min\ & q(x) = (x_1 - 1)^2 + (x_2 - 2.5)^2 \\
s.t.\ & x_1 - 2x_2 + 2 \geq 0 \\
& -x_1 - 2x_2 + 6 \geq 0 \\
& -x_1 + 2x_2 + 2 \geq 0 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
$$

Iteration k	x_k	W_k	p_k	α_k
0	*(2,0)*	$\{3,5\}$	*(0,0)*	
1	*(2,0)*	$\{5\}$	*(-1,0)*	*1*
2	*(1,0)*	$\{5\}$	*(0,0)*	
3	*(1,0)*	$\{\}$	*(0,2.5)*	*0.6*
4	*(1, 1.5)*	$\{1\}$	*(0.4, 0.2)*	*1*
5	*(1.4, 1.7)*	$\{1\}$	*(0,0)*	

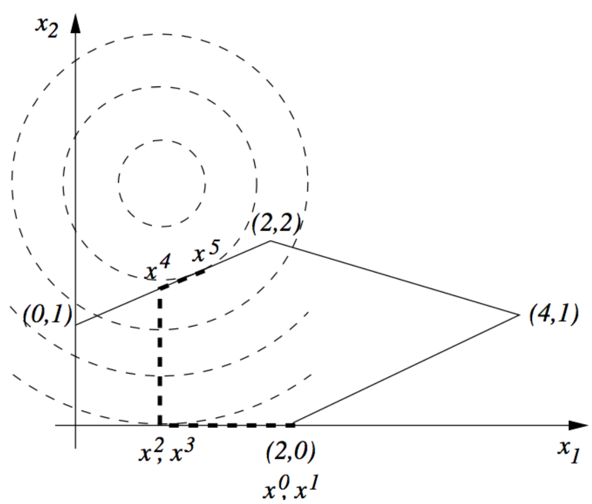

Figure 1: Iterate points of the active set method

References

[1] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, New York, 2006. Second Edition.