Example 16.4 – in more detail

This note aims to go through Example 16.4 in Nocedal & Wright (2006) in some more detail.

Example 16.4 use Algorithm 16.3 (Active Set Method for Convex QP) to solve the following 2-dimensional convex QP

$$\min_x q(x) = (x_1 - 1)^2 + (x_2 - 2.5)^2$$

$$\text{subject to} \quad x_1 - 2x_2 + 2 \geq 0 \tag{1}$$

$$-x_1 - 2x_2 + 6 \geq 0 \tag{2}$$

$$-x_1 + 2x_2 + 2 \geq 0 \tag{3}$$

$$x_1 \geq 0 \tag{4}$$

$$x_2 \geq 0 \tag{5}$$

The constraints are referred to by indices from (1) to (5). The matrices (and vectors) defining the QP problem, are

$$G = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad c = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$$

$$a_1 = \begin{bmatrix} 1 & -2 \end{bmatrix}^\mathsf{T}, \quad b_1 = -2$$

$$a_2 = \begin{bmatrix} -1 & -2 \end{bmatrix}^\mathsf{T}, \quad b_2 = -6$$

$$a_3 = \begin{bmatrix} -1 & 2 \end{bmatrix}^\mathsf{T}, \quad b_3 = -2$$

$$a_4 = \begin{bmatrix} 1 & 0 \end{bmatrix}^\mathsf{T}, \quad b_4 = 0$$

$$a_5 = \begin{bmatrix} 0 & 1 \end{bmatrix}^\mathsf{T}, \quad b_5 = 0$$

Due to the simplicity of the problem, an initial feasible point is easy to determine, and we choose $x^0 = (2, 0)^\mathsf{T}$ (as in the book). At this point (see Figure 1), constraints 3 and 5 are active, and we choose $\mathcal{W}_0 = \{3, 5\}$ (still as in the book – we could just as well have chosen $\mathcal{W}_0 = \{3\}$, $\mathcal{W}_0 = \{5\}$ or $\mathcal{W}_0 = \emptyset$). The iterates on x and p will be denoted with superscript, x^k and p^k, to avoid mix-up with the elements of the vectors x and p.

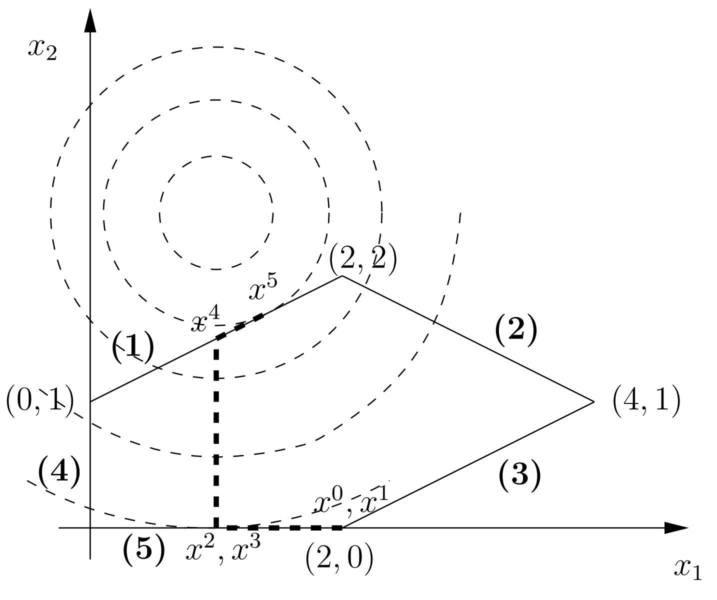

Figure 1: Feasible set, objective function and iterates of the active set method

We then start Algorithm 16.3:

$k = 0$: The EQP (16.39) becomes

$$\begin{aligned} \min_p \quad & \frac{1}{2} p^\mathsf{T} G p + g_0^\mathsf{T} p \\ \text{subject to} \quad & a_3^\mathsf{T} p = -p_1 + 2p_2 = 0 \\ & a_5^\mathsf{T} p = p_2 = 0. \end{aligned}$$

Since we have two (linearly independent) equality constraints, there are no degrees of freedom to optimize, thus the equality constraints completely decide the solution: $p^0 = (0,0)^\mathsf{T}$.

The next step is to use (16.42) to calculate the Lagrange multipliers:

$$a_3\lambda_3 + a_5\lambda_5 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \lambda_3 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \lambda_5 = g_0 = Gx^0 + c = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$$

giving $\lambda_3 = -2$ and $\lambda_5 = -1$. Since not all Lagrange multipliers are positive, the KKT conditions are not fulfilled, and we cannot stop yet. We continue by removing the constraint corresponding to the most negative multiplier (in this case 3). We thus set $\mathcal{W}_1 = \mathcal{W}_0 \backslash \{3\} = \{5\}$, and $x^1 = x^0 = (2,0)^\mathsf{T}$.

$k = 1$: Since we have removed a constraint, the EQP now has one degree of freedom:

$$\begin{aligned} \min_p \quad & \frac{1}{2} p^\mathsf{T} G p + g_1^\mathsf{T} p \\ \text{subject to} \quad & a_5^\mathsf{T} p = p_2 = 0 \end{aligned}$$

We could solve this by solving the full KKT system, but due to the simplicity of the constraint, we here use the reduced-space method by inserting $p_2 = 0$ into the objective function:

$$\min_p \quad \frac{1}{2} \begin{bmatrix} p_1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} p_1 \\ 0 \end{bmatrix} + \left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} \right)^\mathsf{T} \begin{bmatrix} p_1 \\ 0 \end{bmatrix} = p_1^2 - 2p_1.$$

Since this is convex and unconstrained, the minimum is given by setting the derivate equal to zero, giving $p_1 = -1$. Thus $p^1 = (-1,0)^\mathsf{T} \neq 0$.

We use (16.41) to compute α: Of the indexes not in $\mathcal{W}_1$, we have

$$a_1^\mathsf{T} p^1 = -1, \quad a_2^\mathsf{T} p^1 = 1, \quad a_3^\mathsf{T} p^1 = 1, \quad a_4^\mathsf{T} p^1 = 0.$$

The only negative is for constraint 1, for which we calculate

$$\frac{b_1 - a_1^\mathsf{T} x^1}{a_1^\mathsf{T} p^1} = \frac{-2-2}{-1} = 4.$$

Thus, $\alpha_1 = \min(1,4) = 1$, and there are no blocking constraints, and we can set $\mathcal{W}_2 = \mathcal{W}_1 = \{5\}$, and $x^2 = x^1 + 1 \cdot p^1 = (1,0)^\mathsf{T}$.

$k = 2$: Since $\mathcal{W}_2 = \mathcal{W}_1$, the constraints to the EQP remain the same, and we continue to use the reduced space method:

$$\min_p \quad \frac{1}{2} \begin{bmatrix} p_1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} p_1 \\ 0 \end{bmatrix} + \left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} \right)^\mathsf{T} \begin{bmatrix} p_1 \\ 0 \end{bmatrix} = p_1^2.$$

The minimum is obviously $p_1 = 0$, giving $p^2 = (0,0)^\mathsf{T}$. We continue by calculating λ_5,

$$a_5\lambda_5 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \lambda_5 = g_2 = Gx^2 + c = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ -5 \end{bmatrix},$$

giving $\lambda_5 = -5$. We remove this constraint, $\mathcal{W}_3 = \mathcal{W}_2 \backslash \{5\} = \emptyset$, and set $x^3 = x^2$.

$k = 3$: Since there are now no active constraints, we find the solution to

$$\min_p \quad \frac{1}{2} p^\mathsf{T} G p + g_2^\mathsf{T} p$$

by differentiating and setting the derivative equal to zero:

$$Gp + g_2 = 0.$$

We solve this linear equation by inverting G (trivial in this case):

$$p^3 = -G^{-1} g_2 = -\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 2.5 \end{bmatrix}.$$

Since $p^3 \neq 0$, we must first check if there are blocking constraints. We have

$$a_1^\mathsf{T} p^3 = -5, \quad a_2^\mathsf{T} p^3 = -5, \quad a_3^\mathsf{T} p^3 = 5, \quad a_4^\mathsf{T} p^3 = 0, \quad a_5^\mathsf{T} p^3 = 2.5.$$

The two first are negative, and we calculate

$$\frac{b_1 - a_1^\mathsf{T} x^3}{a_1^\mathsf{T} p^3} = \frac{-2-1}{-5} = 0.6 \quad \text{and} \quad \frac{b_2 - a_2^\mathsf{T} x^3}{a_2^\mathsf{T} p^3} = \frac{-6-1}{-5} = 1.4.$$

Now, we see that $\alpha_3 = \min(1, 0.6, 1.4) = 0.6$, and the blocking constraint is the first constraint (see Figure 1). We get $x^4 = x^3 + 0.6 p^3 = (1, 1.5)^\mathsf{T}$, and $\mathcal{W}_4 = \mathcal{W}_3 + \{1\} = \{1\}$.

$k = 4$: The EQP is now

$$\begin{aligned} \min_p \quad & \frac{1}{2} p^\mathsf{T} G p + g_4^\mathsf{T} p \\ \text{subject to} \quad & a_1^\mathsf{T} p = p_1 - 2p_2 = 0. \end{aligned}$$

Although we could still use the reduced space method (but now we would have to calculate the matrices Z and Y), we use now the full space method, that is, we solve the EQP by solving the linear equation system

$$\begin{bmatrix} G & a_1^\mathsf{T} \\ a_1 & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda \end{bmatrix} = \begin{bmatrix} g_4 = Gx^4 + c \\ h = a_1^\mathsf{T} x^4 - b_1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -2 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} \\ -2 - (-2) \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix}$$

We solve this by Matlab[1]

```
>> [2,0,1;0,2,-2;1,-2,0]\[0,-2,0]'

ans =

   -0.4000
   -0.2000
    0.8000
```

and see that $p^4 = (0.4, 0.2)^\mathsf{T} \neq 0$. We have

$$a_2^\mathsf{T} p^4 = -0.8, \quad a_3^\mathsf{T} p^4 = 0, \quad a_4^\mathsf{T} p^4 = 0.4, \quad a_5^\mathsf{T} p^4 = 0.2.$$

[1] Here we use the Matlab 'left matrix divide', see help mldivide. This function checks the structure of the matrix and solves using the most efficient method (in this case most likely using LDL-decomposition, since the KKT matrix is symmetric put not positive (or negative) definite). Using decomposition methods is much faster and more robust than solving the system by calculating the inverse. For a general active set QP solver, it is crucial (for efficiency) to use as fast linear algebra methods as possible.

Only the first is negative, and we calculate

$$\frac{b_2 - a_2^\top x^4}{a_2^\top p^4} = \frac{-6-(-4)}{-0.8} = 2.5.$$

(Note the significance of this: From x^4, we can travel along p^4 2.5 times the length of p^4 before we hit a constraint.) Hence, $\alpha_5 = \min(1, 2.5) = 1$, and $x^5 = x^4 + 1 \cdot p^4 = (1.4, 1.7)^\top$. There are no blocking constraints, thus $\mathcal{W}_5 = \mathcal{W}_4 = \{1\}$.

$k = 5$: The EPQ becomes (similarly to last iteration)

$$\begin{bmatrix} G & a_1^\top \\ a_1 & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda \end{bmatrix} = \begin{bmatrix} g_4 = Gx^4 + c \\ h = a_1^\top x^4 - b_1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & -2 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} -p \\ \lambda \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1.4 \\ 1.7 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} \\ -2-(-2) \end{bmatrix} = \begin{bmatrix} 0.8 \\ -1.6 \\ 0 \end{bmatrix}$$

We againg solve this by Matlab

```
>> [2,0,1;0,2,-2;1,-2,0]\[0.8,-1.6,0]'

ans =

         0
         0
    0.8000
```

and thus $p^5 = (0, 0)^\top$. Now, by solving the EQP by the full space method, we have already calculated the Lagrange multiplier, but we do it again:

$$a_1 \lambda_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \lambda_1 = g_5 = Gx^5 + c = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1.4 \\ 1.7 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \end{bmatrix} = \begin{bmatrix} 0.8 \\ -1.6 \end{bmatrix}$$

to find $\lambda_1 = 0.8$, which is positive, and thus $x^* = x^5 = (1.4, 1.7)^\top$ is the solution.