Scientific Computing 2

Summer term 2017
Prof. Dr. Ira Neitzel
Christopher Kacwin

universitätbonn

Sheet 7

Submission on **Thursday, 22.6.2017**.

Exercise 1. (active set strategy)

We consider the optimization problem

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}(x-x_d)^\top H(x-x_d)$$

$$a_i \le x_i \le b_i\,, \quad i = 1,\dots,n$$

where $H \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, $x_d \in \mathbb{R}^n$ and $a, b \in \mathbb{R}^n$ satisfy $a_i < b_i$ for $i = 1,\dots,n$.

a) Calculate the solution using

$$H = \begin{pmatrix} 20 & 1 \\ 1 & 2 \end{pmatrix}, \; x_d = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \; a = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \; b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

b) Show that the solution is the projection of x_d onto the feasible set with respect to the H-scalar product $(x,y)_H = x^\top H y$. State the variational inequality for the solution explicitly.

A useful method for quadradic optimization problems is the *active set strategy*. For the general problem

$$\min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}x^\top H x - c^\top x$$

$$a_i \le x_i \le b_i\,, \quad i = 1,\dots,n$$

we define the active set as

$$\mathcal{A}(x) = \{i : x_i = a_i \text{ or } x_i = b_i\}\,.$$

Furthermore, denote by C the feasible set and P_C the usual projection operator (w.r.t. the euclidean scalar product). Then, the procedure is as follows:

0. Choose a starting point x^0 and set $k = 0$. Compute the active set $\mathcal{A}(x^k)$. While $k = 0$ or $\mathcal{A}(x^{k-1}) \neq \mathcal{A}(x^k)$, do steps 1., 2. and 3.

1. Compute y^{k+1} as the solution to

$$\min_{y\in\mathbb{R}^n} f(y) = \frac{1}{2}y^\top H y - c^\top y$$

$$y_i = x_i^k\,, \quad i \in \mathcal{A}(x^k)\,.$$

2. Do a projection gradient step

$$x^{k+1} = P_C(y^{k+1} - \rho\nabla f(y^{k+1}))\,.$$

3. Raise $k = k + 1$. Compute the active set $\mathcal{A}(x^k)$.

This algorithm 'approximates' the active set and then calculates the solution. The parameter ρ is of minor importance.

c) Solve a) again, using the active set strategy. Do all intermediate steps. Use $x^0 = (0,0)^\top$ as starting point. (Hint: the algorithm should find the solution after 2 iterations.)

d) Let x^* be the solution to the problem

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^\top H x - c^\top x$$

$$a_i \le x_i \le b_i \,, \quad i = 1, \dots, n$$

and assume that we know $\mathcal{A}(x^*)$ and x_i^* for $i \in \mathcal{A}(x^*)$. Let y be the solution to

$$\min_{y \in \mathbb{R}^n} f(y) = \frac{1}{2} y^\top H y - c^\top y$$

$$y_i = x_i^* \,, \quad i \in \mathcal{A}(x^*)\,.$$

Show that $y \in C$, i.e. $a \le y \le b$. Show that $y = x^*$.

e) Show that after step 1. of the algorithm, one has $[\nabla f(y^{k+1})]_i = 0$ for inactive indices i.

f) Show that if one has $\mathcal{A}(x^k) = \mathcal{A}(x^{k-1})$ during the algorithm, then x^k is the solution to the optimization problem.

(20 points)

Programming exercise 1. (penalty method)

The goal of this programming exercise is to implement a simple penalty strategy for optimization problems with inequality constraints.

a) Show that the function $(t)_+^2 = (\max\{0, t\})^2$ is continuously differentiable and compute its derivative.

b) Implement the following penalty method:

0. Choose $x^0 \in \mathbb{R}^n$, $\alpha_0 > 0$, $\beta > 1$ and $\epsilon > 0$. Set $k = 1$. While $\sum_{i=1}^m (g_i(x^{k-1}))_+^2 > \epsilon^2$ or $k = 1$ holds true, do steps 1. and 2.
1. Compute x^k as the solution to the global optimization problem

$$\min_{x \in \mathbb{R}^n} P_k(x) = f(x) + \frac{\alpha_{k-1}}{2} \sum_{i=1}^m (g_i(x))_+^2 \,.$$

Use the gradient descent method with line search to solve this problem, with x^{k-1} as starting point and $\|y_k - y_{k+1}\| < \epsilon$ as stopping criterion (where y_k are the gradient descent iterates).

2. Set $\alpha_k = \beta \alpha_{k-1}$ and raise $k = k + 1$.

c) Test your implementation for $\alpha_0 = 1$, $\beta = 2$, $\epsilon = 10^{-6}$ and the functions

$$f(x) = 1000 - x_1^2 - 2x_2^2 - x_3^2 - x_1 x_2 - x_1 x_3, \quad g(x) = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 - 25 \\ 8x_1 + 14x_2 + 7x_3 - 56 \end{pmatrix}$$

with starting point $x^0 = (3, 0.2, 3)^\top$.

d) Solve the optimization problem and find all KKT-points. The calculated solution is *not* approximating the global solution. Why is this not a contradiction to the theory?

(20 points)

The programming exercise should be handed in either before/after the exercise class on 22.6.17 (bring your own laptop!) or in the HRZ-CIP-Pool, after making an appointment at "kacwin@ins.uni-bonn.de". All group members need to attend the presentation of your solution. Closing date for the programming exercise is the 22.6.2017.