Exam

Numerical Methods and Optimization course
University of Pisa, 2017-10-30

You may use Matlab, pencil or paper, or a calculator (unless explicitly stated in the exercise). You may use the quick reference sheet on Matlab syntax posted on the web page of the course.

Exercise 1. Consider the linear system of equations

$$M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}, \quad \text{where } M = \begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix}$$

and $A, B \in \mathbb{R}^{n \times n}$, $c, d, x, y \in \mathbb{R}^n$.

a. Show that the solution of this linear system is given by

$$y = A^{-1}c, \quad x = (A^T)^{-1}(d - By). \tag{1}$$

b. Show that

$$M^{-1} = \begin{bmatrix} -(A^T)^{-1}BA^{-1} & (A^T)^{-1} \\ A^{-1} & 0 \end{bmatrix} \tag{2}$$

Hint: you do not need to compute M^{-1} from scratch to do it.

c. Write a Matlab `function [x,y] = blocksystem(A,B,c,d)` that computes the solution of the linear system using (**??**) and a QR factorization of A. Make sure that the same factorization is reused to solve both linear systems.

d. Test your function with the following data:

$$A = 10^{-5} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}.$$

What are the returned values of x and y? What is the norm of the residual $\left\| \begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} c \\ d \end{bmatrix} \right\|$?

e. What are the computational costs (as functions of n) of:

   i. Solving the system using the function `blocksystem` above;
   ii. Solving the system using a QR factorization of M (without special optimizations based on the structure of this matrix).

   You may ignore lower-order terms in n.

Exercise 2. Consider the following constrained optimization problem:

$$\begin{cases} \min \ x_1 + x_2 \\ x_1^2 + x_2^2 - 8x_1 - 8x_2 + 31 \leq 0 \\ 4 - x_1 \leq 0 \end{cases}$$

a. Do global optimal solutions exist? Why?

b. Is it a convex problem? Why?

c. Does the constraints qualification hold in any feasible point? Why?

d. Is $x = (4, 5)$ a local minimum? Why?

e. Find all the solutions of the KKT system.

f. Find local minima and global minima.

g. Find the objective function and constraints of the Lagrangian dual problem.

h. Is $\lambda = (1/8, 0)$ an optimal solution of the Lagrangian dual problem? Why?

Exercise 3. Consider the following unconstrained optimization problem:

$$\begin{cases} \min \ x_1^4 + x_2^4 + 2x_3^4 + x_1^2 + x_2^2 + x_3^2 + x_1 x_2 + x_2 x_3 - 5x_1 - 3x_2 - x_3 \\ x \in \mathbb{R}^3 \end{cases}$$

a. Is it a convex problem? Why?

b. Do global minima exist? Why?

c. Is the global minimum unique? Why?

d. Is $x = (0, 0, 0)$ a global minimum? Why?

e. Solve the problem by means of the gradient method with inexact line search setting $\alpha = 0.1$, $\gamma = 0.9$, $\bar{t} = 1$, starting from the point $(10, 10, 10)$ and using $\|\nabla f(x)\| < 10^{-6}$ as stopping criterion. What is the global minimum? How many iterations are needed?

f. Solve the problem by means of the Newton method with inexact line search setting $\alpha = 0.1$, $\gamma = 0.9$, $\bar{t} = 1$, starting from the point $(10, 10, 10)$ and using $\|\nabla f(x)\| < 10^{-6}$ as stopping criterion. What is the global minimum? How many iterations are needed?

Solutions

Exercise 1.

a. Expanding products, we get

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ay \\ A^T x + By \end{bmatrix}.$$

Hence one gets $c = A^{-1}y$ and $A^T x = d - By$, from which we can solve $x = (A^T)^{-1}(d - By)$.

b. It is sufficient to check that $MM^{-1} = I$. Setting $C = -(A^T)^{-1}BA^{-1}$ for brevity:

$$\begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix} \begin{bmatrix} C & (A^T)^{-1} \\ A^{-1} & 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot C + A \cdot A^{-1} & 0 \cdot (A^T)^{-1} + A \cdot 0 \\ A^T \cdot C + B \cdot A^{-1} & A^T \cdot (A^T)^{-1} + B \cdot 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ A^T C + BA^{-1} & I \end{bmatrix},$$

furthermore $A^T C + BA^{-1} = -A^T(A^T)^{-1}BA^{-1} + BA^{-1} = -BA^{-1} + BA^{-1} = 0$.

As an alternative, one can get an expression of M^{-1} by considering (**??**) and rewriting it as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (A^T)^{-1}d - (A^T)^{-1}By \\ A^{-1}c \end{bmatrix} = \begin{bmatrix} (A^T)^{-1}d - (A^T)^{-1}BA^{-1}c \\ A^{-1}c \end{bmatrix} = \begin{bmatrix} -(A^T)^{-1}BA^{-1} & (A^T)^{-1} \\ A^{-1} & 0 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}.$$

c. We need to solve the linear system $Ay = c$ with $A = QR$, using the formula $y = (QR)^{-1}c = R^{-1}Q^T c$, and then the linear system $A^T x = d - By$, using $x = (A^T)^{-1}(d - By) = (R^T Q^T)^{-1}(d - By) = Q(R^T)^{-1}(d - By)$.

```
function [x, y] = blocksystem(A, B, c, d);
[Q, R] = qr(A);
y = R \ (Q'*c);
x = Q * (R' \ (d - B*y));
```

Note that the system with R^T is solved using back-substitution (since R^T is a lower triangular system), and that the parentheses are necessary so that the last two rows have cost $O(n^2)$.

d.

```
>> A = 1e-5*[2 1; 1 2];
>> B = [3 2; 2 1];
>> c = [3;4]; d = [5;6];
>> [x, y] = blocksystem(A, B, c, d)
x =
   1.0e+10 *
   -2.5555
   -0.2222
y =
   1.0e+05 *
    0.6667
    1.6667
>> norm([zeros(2) A; A' B]*[x;y] - [c;d])
ans =
   1.3016e-10
```

e. i. The line `[Q, R] = qr(A);` requires $\frac{4}{3}n^3 + O(n^2)$ floating point operations. The following lines contain only matrix-vector products and the solution of triangular linear systems, both of which cost $O(n^2)$. Hence the total cost is $\frac{4}{3}n^3 + O(n^2)$ flops.

ii. Computing the QR factorization of M costs $\frac{4}{3}(2n)^3 + O(n^2) = \frac{32}{3}n^3 + O(n^2)$ flops, since M has size $2n$. All the other required operations still have quadratic cost.

Exercise 2.

a. Yes, the objective function is continuous and the feasible region is closed and bounded (Weierstrass Theorem).

b. Yes, the objective function and the constraints are convex.

c. Yes, the Slater constraints qualification holds.

d. No, it does not solve the KKT system.

e. The solution of the KKT system is:

$x = (4, 3) \quad \lambda = (1/2, 1)$

f. The global minimum is $x = (4, 3)$.

g. The Lagrangian dual problem is

$$\begin{cases} \max \ \dfrac{-\lambda_2^2 + 2\lambda_2 - 2}{4\lambda_1} - \lambda_1 + 8 \\ \lambda_1 > 0 \\ \lambda_2 \geq 0 \end{cases}$$

h. No. The optimal value of the dual is 7, while $\varphi(1/8, 0) = 31/8 < 7$.

Exercise 3.

a. Yes, the hessian matrix of the objective function f is

$$\nabla^2 f(x) = \begin{pmatrix} 12x_1^2 + 2 & 1 & 0 \\ 1 & 12x_1^2 + 2 & 1 \\ 0 & 1 & 24x_3^2 + 2 \end{pmatrix}$$

Notice that $\nabla^2 f(x) = D(x) + Q$, where

$$D(x) = \begin{pmatrix} 12x_1^2 & 0 & 0 \\ 0 & 12x_1^2 & 0 \\ 0 & 0 & 24x_3^2 \end{pmatrix} \qquad Q = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

$D(x)$ is positive semidefinite for any $x \in \mathbb{R}^3$ and Q is positive definite, hence f is strongly convex.

b. Yes, f is strongly convex and the feasible region is closed and convex.

c. Yes, f is strongly convex and the feasible region is closed and convex.

d. No, $(0, 0, 0)$ is not a stationary point.

e. After 52 iterations the algorithm finds the approximated global minimum

$$x = (0.8741, 0.5801, 0.1847).$$

f. After 24 iterations the algorithm finds the approximated global minimum

$$x = (0.8741, 0.5801, 0.1847).$$