Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180
Optimization Theory
Spring 2016

Exercise set 3 solution

1 In the following, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. We also define

$$f(x) = ||Ax - b||^2 = x^T A^T Ax - 2x^T A^T b + b^T b$$

and observe that

$$\nabla f(x) = 2A^T Ax - 2A^T b,$$
$$\nabla^2 f(x) = 2A^T A$$

a) We wish to show that $x^* \in \mathbb{R}^n$ solves the least squares problem

$$\min_{x\in\mathbb{R}^n} f(x) \tag{1}$$

if and only if x^* satisfies the normal equations

$$A^T Ax^* = A^T b.$$

First, if x^* solves (1), then x^* is a global minimizer of f, so $\nabla f(x^*) = 0$, meaning $A^T Ax^* = A^T b$.

Conversely, if $A^T Ax^* = A^T b$, then $\nabla f(x^*) = 0$. We may note that f is convex, since $\nabla^2 f(x)$ is symmetric positive semi-definite. Since f is also differentiable, any stationary point is a global minimizer, by Theorem 2.5 in N&W. Hence, x^* is a global minimizer of f, that is, x^* solves (1).

b) We wish to show that the optimization problem (1) admits a solution $x^* \in \mathbb{R}^n$. By the previous problem, this is equivalent to checking whether the equation

$$A^T Ax^* = A^T b$$

has a solution. This will be the case if $A^T b \in \operatorname{ran} A^T A$. First, we may observe that clearly, $A^T b \in \operatorname{ran} A^T$. Next, we use the Fundamental Theorem of Linear Algebra and find

$$\operatorname{ran} A^T A = (\ker(A^T A)^T)^\perp = (\ker A^T A)^\perp = (\ker A)^\perp = \operatorname{ran} A^T.$$

Thus, we know that $A^T b \in \operatorname{ran} A^T A$, and the problem admits a solution. In the above, we used that $\ker A^T A = \ker A$. This can be seen from the fact that

$$x \in \ker A \Rightarrow Ax = 0 \Rightarrow A^T Ax = 0 \Rightarrow x \in \ker A^T A$$

and

$$x \in \ker A^T A \Rightarrow A^T Ax = 0 \Rightarrow x^T A^T Ax = 0 \Rightarrow ||Ax|| = 0 \Rightarrow Ax = 0 \Rightarrow x \in \ker A.$$

c) We wish to show that the solution x^* of (1) is unique if the rank of A equals n. If rank $A = n$, then by the Rank-Nullity theorem, dim ker$A = 0$, so ker$A = \{0\}$. This means that $\nabla^2 f(x)$ is SPD, since

$$x^T \nabla^2 f(x) x = x^T A^T A x = ||Ax||^2 \geq 0,$$

with equality if and only if $x = 0$. Thereby, f is strictly convex and has at most one minimizer (see exercise set 1). Since we have already proven that a minimizer must exist, this means that the solution x^* of (1) is unique.

d) We wish to show that, regardless of the rank of A, the optimization problem

$$\min_{x \in \mathbb{R}^n} ||x||^2 \qquad \text{s.t. } x \text{ solves (1)} \tag{2}$$

admits a unique solution $x^\dagger \in \mathbb{R}^n$. First, let us rephrase the problem. Define $g(x) = ||x||^2$, and let $\Omega = \{x \in \mathbb{R}^n \text{ s.t. } x \text{ solves (1)}\}$. We can now equivalently state the problem as

$$\min_{x \in \Omega} g(x)$$

We know that f is a convex function, and therefore Ω is a convex set (see exercise set 1). Furthermore, since $\nabla^2 g(x) = 2I$, which is an SPD matrix, we know that g is a strictly convex function. Also, g is coercive. Therefore, g has a unique minimizer $x^\dagger$.

e) We wish to show that the solution $x^\dagger$ of (2) is uniquely characterized by the conditions $A^T A x^\dagger = A^T b$ and $x^\dagger \in \text{ran}A^T$.

First, we know that problem (2) only considers $x \in \mathbb{R}^n$ such that x solves (1). Thus, any solution of (2) must have the property that $A^T A x^\dagger = A^T b$.

Secondly, we can see that any $x \in \text{ran}A^T$ that solves (1) must be unique. If we assume the opposite and take x and y as two distinct solutions, then

$$A^T A(x - y) = A^T b - A^T b = 0,$$

so either $x = y$ or $x - y \in \ker A^T A$, meaning at least one of x and y is not in $\text{ran}A^T$ since $\ker A^T A = \ker A = (\text{ran}A^T)^\perp$. Either way, we have a contradiction, so any $x \in \text{ran}A^T$ that solves (1) must be unique.

Finally, we identify $x^\dagger$ with this unique element of $\text{ran}A^T$, and show that it is the solution of (2). As a consequence of the above discussion, we see that any solution x of (1) can be written as $x = x^\dagger + y$, where $y \in (\text{ran}A^T)^\perp$. Thus, (2) can be rewritten as

$$\min_{z \in (\text{ran}A^T)^\perp} ||x^\dagger + z||^2.$$

But, since $z \in (\text{ran}A^T)^\perp$ and $x^\dagger \in \text{ran}A^T$, we have $z^T x^\dagger = 0$ and thus

$$||x^\dagger + z||^2 = ||x^\dagger||^2 + ||z||^2,$$

which clearly has a minimum for $z = 0$. Therefore, $x^\dagger$ which is uniquely characterized by $A^T A x^\dagger = A^T b$ and $x^\dagger \in \text{ran}A^T$, is indeed the solution of (2).

2 **a)** As seen in problem 1c), if A has full rank then A^TA is positive definite, and so we know that the CG algorithm applied to the system $A^TAx = A^Tb$ will converge. We will now use induction to show that the algorithm proposed in the exercise is identical to the CG algorithm in the sense that $x_k^A = x_k^{CG}$, where x_k^A denotes iterates from the algorithm in the problem and x_k^{CG} iterates from the CG algorithm.

Our induction claim is fairly involved; we claim that, for any k, the following statements hold:

$$r_{k-1}^{CG} = s_{k-1}^A, \qquad p_{k-1}^{CG} = p_{k-1}^A, \qquad \alpha_{k-1}^{CG} = \alpha_{k-1}^A, \qquad x_k^{CG} = x_k^A. \tag{3}$$

We start by considering the base case $k = 1$, and see that

$$\begin{aligned}
r_0^{CG} &= A^TAx_0 - A^Tb = A^T(Ax_0 - b) = A^Tr_0^A = s_0^A, \\
p_0^{CG} &= -r_0^{CG} = -s_0^A = p_0^A, \\
\alpha_0^{CG} &= \frac{(r_0^{CG})^Tr_0^{CG}}{(p_0^{CG})^TA^TAp_0^{CG}} = \frac{(s_0^A)^Ts_0^A}{(p_0^A)^TA^TAp_0^A} = \alpha_0^A, \\
x_1^{CG} &= x_0 + \alpha_0^{CG}p_0^{CG} = x_0 + \alpha_0^Ap_0^A = x_1^A.
\end{aligned}$$

Next, assuming that the hypotheses (3) hold, we need to show that they also hold for the next step, i.e. that

$$r_k^{CG} = s_k^A, \qquad p_k^{CG} = p_k^A, \qquad \alpha_k^{CG} = \alpha_k^A, \qquad x_{k+1}^{CG} = x_{k+1}^A.$$

This is shown in the following equations:

$$\begin{aligned}
r_k^{CG} &= r_{k-1}^{CG} + \alpha_{k-1}^{CG}A^TAp_{k-1}^{CG} = A^T(r_{k-1}^A + \alpha_{k-1}^AAp_{k-1}^A) = s_k^A, \\
p_k^{CG} &= -r_k^{CG} + \frac{(r_k^{CG})^Tr_k^{CG}}{(r_{k-1}^{CG})^Tr_{k-1}^{CG}}p_{k-1}^{CG} = -s_k^A + \frac{(s_k^A)^Ts_k^A}{(s_{k-1}^A)^Ts_{k-1}^A}p_{k-1}^A = p_k^A, \\
\alpha_k^{CG} &= \frac{(r_k^{CG})^Tr_k^{CG}}{(p_k^{CG})^TA^TAp_k^{CG}} = \frac{(s_k^A)^Ts_k^A}{(p_k^A)^TA^TAp_k^A} = \alpha_k^A, \\
x_{k+1}^{CG} &= x_k^{CG} + \alpha_k^{CG}p_k^{CG} = x_k^A + \alpha_k^Ap_k^A = x_{k+1}^A.
\end{aligned}$$

Thus, the iterates coincide and the algorithms are equivalent.

b) Assuming that A has rank $r < n$ and that $x_0 = 0$, we wish to show that the algorithm in the exercise converges to the solution $x^\dagger$ of the optimization problem (2) in at most r steps.

Two key points here are the relations shown in exercise 1e) and 2a). In 1e), we showed that the solution of (2) is uniquely characterized by the conditions

$$A^TAx^\dagger = A^Tb \quad \text{and} \quad x^\dagger \in \operatorname{ran} A^T. \tag{4}$$

In 2a), we showed that the algorithm in the exercise is equivalent to the CG algorithm applied to the problem $A^TAx = A^Tb$. Therefore, if the algorithm converges, it converges to a solution x^* of the problem $A^TAx = A^Tb$. This x^* therefore satisfies the first condition in (4).

Furthermore, the choice $x_0 = 0$ means that $p_0 = A^T b \in \text{ran}A^T$, such that $x_1 = \alpha_0 p_0 \in \text{ran}A^T$. In addition, all subsequent p_k are of the form

$$p_k = -A^T r_k + \beta_k p_{k-1}.$$

Hence, we see that all $p_k \in \text{ran}A^T$ and thus all subsequent x_k (which are of the form $x_k = x_{k-1} + \alpha_{k-1} p_{k-1}$) also lie in $\text{ran}A^T$. Therefore, if the algorithm converges, it will converge to an $x^\dagger \in \text{ran}A^T$ satisfying both conditions (4).

What remains is then to show that the algorithm does converge, and that it does so using at most r steps. First, we need to check whether all operations in the algorithm are legal; especially, whether we risk division by zero anywhere. The two possibilities for this are in calculating

$$\frac{||s_k||^2}{p_k^T A^T A p_k} \quad \text{and} \quad \frac{||s_{k+1}||^2}{||s_k||^2}.$$

We may observe that $A^T A$ is SPD on $\text{ran}A^T$, since $p_k^T A^T A p_k = ||Ap_k||^2 \geq 0$, with equality if and only if $p_k \in \text{ker}A = (\text{ran}A^T)^\perp$. But, since $p_k \in \text{ran}A^T$ for all k, division by $p_k^T A^T A p_k$ is OK. If either p_k or s_k are zero, we have obtained convergence since, by analogue to the CG algorithm, $s_k = r_k^{CG}$, with $r_k^{CG} = A^T A x_k - A^T b$, and since $p_k \neq 0$ unless $s_k = 0$.

Since all operations in the algorithm are legal, we can use the equivalence of our algorithm and the CG algorithm for the problem $A^T A x = A^T b$ and follow the proof of Theorem 5.3 in N&W, which shows that the p_k generated by the CG algorithm (and thus our algorithm) are conjugate with respect to $A^T A$. The only major difference from the proof of Theorem 5.3 is that at the end of the proof, Theorem 5.1 is invoked, stating that the algorithm will converge in at most n iterations. In our case, this reduces to at most r iterations, since the solution $x^\dagger$ lies in the r-dimensional subspace $\text{ran}A^T \subset \mathbb{R}^n$, spanned by the r linearily independent vectors $p_0, ..., p_{r-1}$.

3 The file `conjGrad.m` on the course webpage contains an implementation of the CG method and additional code for running the required tests, along with comments on the produced figures.

4 The file `conjGrad.m` on the course webpage contains an implementation of the CG method and additional code for running the required tests, along with comments on the produced figures.