University of Vienna
Faculty of Mathematics

Nonlinear Optimization Problems

Milutin Popovic

January 15, 2023

Contents

1 Sheet 6

1.1 Exercise 35

Let $M \in \mathbb{R}^{n\times n}$ with $\|M\| < 1$. Show that $I - M$ is regular and

$$\|\left(I-M\right)\| \leq \frac{1}{1-\|M\|}. \tag{1}$$

Suppose $I - M$ is not singular then for $x \in \mathbb{R}^n$ we have that

$$(I-M)\,x = 0 \tag{2}$$
$$\Leftrightarrow Ix - Mx = 0 \tag{3}$$
$$\Leftrightarrow Mx = x. \tag{4}$$

But since $\|M\| < 1$ then $\forall x \in \mathbb{R}^n$ we have that $\|Mx\| < \|x\|$. This means that

$$\ker\left(I-M\right) = \emptyset, \tag{5}$$

so $I-M$ is regular. The identity on the other hand is derived by the following observation

$$\left(I-M\right)^{-1} - \left(I-M\right)^{-1} M = \left(I-M\right)\left(I-M\right)^{-1} = I, \tag{6}$$

Then we calculate

$$\|\left(I-M\right)^{-1}\| = \|I + \left(I-M\right)^{-1} M\| \tag{7}$$
$$\leq \|I\| + \|(I-M)^{-1}\|\|M\|, \tag{8}$$

rearranging gives

$$\|\left(I-M\right)^{-1}\| - \|(I-M)^{-1}\| \leq \|I\| \tag{9}$$
$$\|(I-M)^{-1}\|\left(1-\|M\|\right) \leq 1 \tag{10}$$
$$\|\left(I-M\right)^{-1}\| \leq \frac{1}{1-\|M\|}. \tag{11}$$

Now let $A, B \in \mathbb{R}^{n\times n}$ with $\|I - BA\| < 1$. Show that A and B are regular and that

$$\|B^{-1}\| \leq \frac{\|A\|}{1 - \|I - BA\|} \tag{12}$$
$$\|A^{-1}\| \leq \frac{\|B\|}{1 - \|I - BA\|} \tag{13}$$
$$\tag{14}$$

We know that for $M \in \mathbb{R}^{n\times n}$ with $\|M\| < 1$ then $I - M$ is regular and the inequality in 1 holds. Set $M = I - BA$ then $I - M = AB$ is regular. Because AB is regular so is A and B. Now note that for all regular matrices we have that $\|A^{-1}\| \leq \|A\|^{-1}$. Furthermore

$$\| (AB)^{-1} \| \leq \|B^{-1}\|\|A^{-1}\|. \tag{15}$$

Then for A we have

$$\|A^{-1}\| \leq \frac{1}{\|B^{-1}\|}\frac{1}{1 - \|I - BA\|} \leq \frac{\|B\|}{1 - \|I - BA\|}. \tag{16}$$

and for B

$$\|B^{-1}\| \leq \frac{1}{\|A^{-1}\|}\frac{1}{1 - \|I - BA\|} \leq \frac{\|A\|}{1 - \|I - BA\|}. \tag{17}$$

1.2 Exercise 36

Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^4 + 2x^2y^2 + y^4$. Show that the local Newton algorithm converges to the unique global minimum of f for every $(x^0, y^0) \in \mathbb{R}^2\backslash\{(0,0)^T\}$. First we determine the minimum x^*. Note that $f(x, y) = \left(x^2 + y^2\right)^2 \geq 0$ for all $(x, y)^T \in \mathbb{R}^2$. Since f is strongly convex the only minimum, which is the global minimum is $(x, y)^T = (0, 0)^T$. The Hessian of f is

$$\nabla^2 f(x, y) = \begin{pmatrix} 12x^2 + 4y^2 & 8xy \\ 8xy & 12y^2 + 4x^2 \end{pmatrix}. \tag{18}$$

Now note that the Hessian at the minimum $\nabla^2 f(0, 0)$ is the zero matrix which is singular. But considering starting vectors $(x, y)^T \neq (0, 0)^T$, all we need in the local Newton algorithm is the solution to the equation $\nabla^2 f(x^k)d_k = -\nabla f(x^k)$. Meaning we need to show that $\nabla^2 f(x, y)$ is regular for all $(x, y)^T \neq (0, 0)^T$, in this case we look at the determinant of the Hessian

$$\det(\nabla^2 f(x, y)) = 48 * (x^2 + y^2)^2 > 0 \ \ \forall (x, y)^T \neq (0, 0)^T. \tag{19}$$

This means that the sequence $(x^k)_{k\geq 0}$ produced by the local newton algorithm converges to the unique global minimum of f given by $(0, 0)^T$. Indeed if we calculate the solution of the system $\nabla^2 f(x, y)d = -\nabla f(x, y)$ we get that $d = (\frac{x}{3}, \frac{y}{3})^T$.

1.3 Exercise 37

Show that the local Newton Algorithm is invariant to affine-linear transformation, for a regular matrix $A \in \mathbb{R}^{n\times n}$ and $c \in \mathbb{R}^n$, $(x^k)_{k\geq 0}$ the sequence generated by the local Newton algorithm for minimizing f with starting vector x^0. Then let $(y^k)_{k\geq 0}$ the sequence generated by the local Newton algorithm for the function $g(y) := f(Ay + c)$ with starting vector y^0, then

$$x^0 = Ay^0 + c \quad \Longrightarrow x^k = Ay^k + c \ \ \forall k \geq 0. \tag{20}$$

First of all we calculate the gradient and the hessian for g

$$\nabla g(y) = \nabla f(Ay + c) = A^T \nabla f(Ay + c) \tag{21}$$
$$\nabla^2 g(y) = \nabla^2 f(Ay + c) = A^T \nabla^2 f(Ay + c)A \tag{22}$$
$$\tag{23}$$

Now we need to prove that $x^{k+1} = Ay^{k+1} + c$

$$x^{k+1}q = x^k + d_k = x^k - \left(\nabla^2 f(x^k)\right)^{-1} \nabla f(x_k) \tag{24}$$
$$= Ay^k + c - \left(\nabla^2 (f(Ay^k + c)\right)^{-1} \nabla f(Ay^k + c) \tag{25}$$
$$= Ay^k + c - AA^{-1} \left(\nabla^2 (f(Ay^k + c)\right)^{-1} \nabla f(Ay^k + c) \tag{26}$$
$$= Ay^k + c - AA^{-1}A^T A^{-T} \left(\nabla^2 (f(Ay^k + c)\right)^{-1} \nabla f(Ay^k + c) \tag{27}$$
$$= A\left(y^k - \left(A^T \nabla^2 f(Ay^k + c)A\right)^{-1} A^T \nabla f(Ay^k + c)\right) + c \tag{28}$$
$$= Ay^{k+1} + c. \tag{29}$$

by that the induction is finished.

1.4 Exercise 38

Let $M \in \mathbb{R}^{n\times n}$ be a regular matrix and $\{M\}_{k\geq 0} \in \mathbb{R}^{n\times n}$ a sequence of matrices which converge to M as $k \to \infty$. Sow that there exists a $k_0 \geq 0$ such that M_k is regular for all $k \geq k_0$ and the sequence $\{M_k^{-1}\}_{k\geq 0}$ converges to M^{-1}.
The map $M \to M^{-1}$ is a continuous invertible meaning it is monotone. $M^{-1} = \frac{\text{adj}(M)}{\det(M)}$. Then convergence means that there is a $k \geq k_0$ such that for all $M_k \in B_{\frac{1}{k}}(M)$ we have that $\|M_k - M\| < \frac{1}{k}$ then M_k is sufficiently close to M and so regular. Since $\{M_k\}_{k\geq k_0} \cup M$ is a compact set of invertible matrices so is $\{M_k^{-1}\}_{k\geq k_0} \cup M^{-1}$, meaning it is bounded. This means that $\{M_k^{-1}\}_{k\geq k_0}$ converges to M^{-1}.

1.5 Exercise 39

Let $H \in \mathbb{R}^{n\times n}$ be regular $u, v \in \mathbb{R}^n$ arbitrary. Show that $H + uv^T$ regular $\Leftrightarrow 1 + v^T H^{-1} u \neq 0$, then the Sherman-Morrison formula holds

$$\left(H + uv^T\right)^{-1} = \left(I - \frac{1}{1 - v^T H^{-1} u} H^{-1} uv^T\right) H^{-1} \tag{30}$$

Let $1 + v^T H^{-1} u = 0$ then

$$\det\left(H + uv^T\right) = (1 + v^T H^{-1} u)\det(H) = 0. \tag{31}$$

This means that H is not invertible. Now we need to check if the inverse really holds which is done by simply multiplying

$$\left(H + uv^T\right)\left(H^{-1} - \frac{H^{-1}uv^T H^{-1}}{1 + v^T H^{-1} u}\right) = \tag{32}$$
$$= HH^{-1} + uv^T H^{-1} - H\frac{H^{-1}uv^T H^{-1}}{1 + v^T H^{-1} u} uv^T \frac{H^{-1}uv^T H^{-1}}{1 + v^T H^{-1} u} \tag{33}$$
$$= I + uv^T H^{-1} - \frac{uv^T H^{-1} + uv^T H^{-1} uv^T H^{-1}}{1 + v^T H^{-1} u} \tag{34}$$
$$= I + uv^T H^{-1} - \frac{u\left(1 + v^T H^{-1} u\right) v^T H^{-1}}{1 + v^T H^{-1} u} \tag{35}$$
$$= I + uv^T H^{-1} - uv^T H^{-1} \tag{36}$$
$$= I \tag{37}$$

Since these are square matrices $AB = I$ is the same as $BA = I$.

1.6 Exercise 40

Consider the quadratic optimization problem

$$\begin{aligned} \min \quad & f(x) := \gamma + c^T x + \frac{1}{2} x^T Q x, \\ \text{s.t} \quad & h(x) := b^T x = 0, \end{aligned} \tag{38}$$

with $Q \in \mathbb{R}^{n\times n}$ SPD , $b, c \in \mathbb{R}^n$, $b \neq 0$ and $\gamma \in \mathbb{R}$. For a given $\alpha > 0$ find the minimum $x^*(\alpha)$ of the penalty function

$$P(x;\alpha) := f(x) + \frac{\alpha}{2}\left(h(x)\right)^2 \tag{39}$$

determine $x^* := \lim_{\alpha\to\infty} x^*(\alpha)$ and prove that x^* is a unique optimal solution of the optimization problem in 38. We start with calculating the minimum of $P(x(\alpha))$

$$\begin{aligned}
\nabla P(x(\alpha)) &= \nabla f(x) + \frac{\alpha}{2}\nabla h(x)^2 && (40)\\
&= c + Qx + \frac{\alpha}{2}2h(x)\nabla h(x) && (41)\\
&= c + Qx + \alpha b^T x b = 0 && (42)
\end{aligned}$$

$$\begin{aligned}
Qx + \alpha b b^T x &= -c && (43)\\
\left(Q + \alpha b b^T\right) x &= -c. && (44)
\end{aligned}$$

Using the Sherman-Morrison formula in 30 for $H = Q$, $u = \alpha b$ and $v = b$ we get

$$x^*(\alpha) = \left(\frac{\alpha}{1 + \alpha b^T Q^{-1} b} Q^{-1} b b^T - I\right) Q^{-1} c \tag{45}$$

The limit is then (the standard limit $\frac{x}{1+kx} \to \frac{1}{k}$ as x goes to infinity)

$$\begin{aligned}
x^* &= \lim_{\alpha\to\infty} x^*(\alpha) && (46)\\
&= \left(\frac{Q^{-1} b b^T}{b^T Q^{-1} b} - I\right) Q^{-1} c. && (47)
\end{aligned}$$

To show that x^* is a unique solution of the optimization problem 38 we need to show that it satisfies the optimality condition and that it is unique. Now it is unique because Q is SPD meaning it is regular and invertible and $\nabla^2 f = Q > 0$. Further more (x^*, α) is a KKT point of $P(x, \alpha)$ then x^* is a minumum of the optimization problem. Now we show that $b^T x^* = 0$:

$$\begin{aligned}
b^T x^* &= \left(\frac{b^T Q^{-1} b b^T Q^{-1}}{b^T Q^{-1} b} - b^T Q^{-1}\right) c && (48)\\
&= \left(b^T Q - b^T Q\right) c && (49)\\
&= 0. && (50)
\end{aligned}$$