Nonlinear Optimization – Sheet 11

Exercise 1

(*i*) This is just the classical equation $z^2 - 1 = 0$ which has the solution set $\{\pm 1\}$. The perturbed generalized equation is of the form

$$\Delta = F(z^*) + F'(z^*)(w - z^*) = \begin{cases} 2(w-1) & z^* = 1 \\ -2(w+1) & z^* = -1 \end{cases}$$

and has the unique solution

$$w = \begin{cases} \Delta/2 + 1 & z^* = 1 \\ -\Delta/2 - 1 & z^* = -1\,, \end{cases}$$

which clearly depends Lipschitz-continuously on Δ. Therefore both solutions are strongly regular.

(*ii*) This generalized equations reduces to the inequality $z^2 - 1 \le 0$, which has the solution set $[-1, 1]$. The perturbed equation

$$\Delta \ge F(z^*) + F'(z^*)(w - z^*) = 2wz^* - (z^*)^2 - 1$$

is not locally uniquely solvable around any $z^* \in [0, 1]$.

(*iii*) First off, note that $\mathcal{N}_{\mathbb{R}_\ge}(z)$ is empty for $z \notin \mathbb{R}_\ge$, therefore the solution set is contained in $\mathbb{R}_\ge$. For $z = 0$ we have $\mathcal{N}_{\mathbb{R}_\ge}(z) = \mathbb{R}_{\le 0}$ and the generalized equation reads as $F(z) \ge 0$. For $z > 0$ we have $\mathcal{N}_{\mathbb{R}_\ge}(z) = 0$ and the generalized equation reads as $F(z) = 0$. Therefore the solution set is $\{0, 1\}$. Strong regularity is omitted.

Exercise 2

See the attached code file.

Exercise 3

Lines 3 and 12 set ζ to be the residual $Ad - b$. Lines 4 and 13 solve

$$\begin{pmatrix} M & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} p \\ 0 \end{pmatrix} = - \begin{pmatrix} \zeta \\ 0 \end{pmatrix}$$

This is the KKT system of the quadratic problem of minimizing $\frac{1}{2} p^t M p + (M^{-1}\zeta)^t p$ such that $Bp = 0$, as seen in equation (13.2). Lemma 11.2 states that the solutions of the KKT system are precisely the global solutions to the problem. Careful inspection shows that this global solution is the projection of $-M^{-1}\zeta$ onto the kernel of B, which is what we want because $-M^{-1}\zeta$ is the steepest descent direction (4.7).

Exercise 4

(i) As M is s.p.d., the function

$$\frac{1}{2} \|p - M^{-1}\nabla f(d^*)\|_M^2 = \frac{1}{2} p^\intercal M p - p^\intercal \nabla f(d^*) + \text{const}$$

is convex in p. Therefore, its minimizer $p^* := \operatorname{proj}^M_{\ker B}(\nabla_M f(d^*))$ satisfies the KKT-conditions, i.e. we find a λ s.t.

$$Mp - \nabla f(d^*) + B^\intercal \lambda = 0.$$

If d^* is a solution to (13.1), then it satisfies the KKT-conditions, i.e. we find a $\tilde{\lambda}$ s.t.

$$\nabla f(d^*) + B^\intercal \tilde{\lambda} = 0.$$

Adding that to the previous equation, we obtain

$$Mp = B^\intercal \overline{\lambda} \iff p = M^{-1} B^\intercal \overline{\lambda}$$

We know

$$0 = Bp = BM^{-1} B^\intercal \overline{\lambda}$$

If $m \le n$, the matrix on the RHS is invertible as the rank of B is m. In that case, $\overline{\lambda} = 0$ and therefore $p = M^{-1} B^\intercal \overline{\lambda} = 0$, too. If, on the other hand, $m > n$, by the dimension formula we obtain $\ker B = 0$ and $Bp = 0$ implies $p = 0$ directly.

(ii) Set $s := d - d^*$ and let π denote the M-orthogonal projection onto $\ker B$. We have $s \in \ker B$ and therefore $\pi s = s$. This yields

$$\begin{aligned}
\|s\|_A^2 &= s^t A s \\
&= s^t M M^{-1} A s \\
&= \langle s, M^{-1} A s \rangle_M \\
&= \langle \pi s, M^{-1} A s \rangle_M \\
&= \langle s, \pi M^{-1} A s \rangle_M \\
\text{(CSU)} \qquad &\le \|s\|_M \cdot \|\pi M^{-1} A s\|_M \,.
\end{aligned}$$

We have

$$M^{-1} A s = M^{-1} A d - M^{-1} A d^* = M^{-1}(Ad - b) - M^{-1}(Ad^* - b) = \nabla_M f(d) - \nabla_M f(d^*)\,.$$

In (i), we showed $\pi \nabla_M f(d^*) = 0$, therefore $\pi M^{-1} A s = \pi \nabla_M f(d)$. For the other factor, find an $t \in \mathbb{R}^{n-m}$ such that $s = Zt$. Then consider

$$\begin{aligned}
\|s\|_M^2 &= s^t M s \\
&= t^t Z^t M Z t \\
(Z \text{ is } M\text{-orthogonal}) \qquad &= t^t t \\
&= t^t (Z^t A Z)^{-3/2} (Z^t A Z)^{1/2} t \\
\text{(CSU)} \qquad &\le \|(Z^t A Z)^{-3/2} t\| \cdot \|(Z^t A Z)^{1/2} t\| \\
(\text{defining ineq. for matrix-norm of } (Z^t A Z)^{-1}) \qquad &\le \|(Z^t A Z)^{-1}\| \cdot \|(Z^t A Z)^{1/2} t\|^2 \\
&= \|(Z^t A Z)^{-1}\| \cdot t^t Z^t A Z t \\
&= \|(Z^t A Z)^{-1}\| \cdot s^t A s \\
&= \|(Z^t A Z)^{-1}\| \cdot \|s\|_A^2 \,.
\end{aligned}$$

This concludes the proof.