Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization I Spring 2020

Solutions to exercise set 8

1 Let

$$f(x) = x_1^4 + 2x_2^4 + x_1x_2 + x_1 - x_2 + 2.$$

Starting at the point $x_0 = (0, 0)$ compute explicitly one step for the trust region method with the model function $m(p) = f(x_0) + g^\top p + \frac{1}{2}p^\top Bp$, where $g = \nabla f(x_0)$, $B = \nabla^2 f(x_0)$, and the trust region radius $\Delta = 1$.

Solution: We invoke Theorem 4.1 in Nocedal & Wright, which says that p_0 is a global minimizser to the trust-region subproblem

$$\min_{\|p\|\leq\Delta} m(p),$$

with $\Delta = 1$, if and only if there exists a $\lambda \geq 0$ such that

$$(B + \lambda\,\mathrm{Id})p_0 = -g, \tag{1}$$

$$\lambda(\Delta - \|p_0\|) = 0, \text{ and} \tag{2}$$

$$B + \lambda\,\mathrm{Id} \quad \text{is positive semi-definite.} \tag{3}$$

Routine calculations yield that

$$g = \nabla f(x_0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \text{and} \qquad B = \nabla^2 f(x_0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Since B has eigenvalues ± 1, we must have $\lambda \geq 1$ in order to guarantee the positive semi-definiteness of the matrix $B + \lambda\,\mathrm{Id}$. As a result, from complementarity condition (2) we must have $\|p_0\| = 1$, so p_0 lies on the trust-region boundary.

Solution of (1) equals

$$p_0 = \frac{1}{1-\lambda}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

provided $\lambda \neq 1$ (there is no solution for $\lambda = 1$), and from the conditions $\|p_0\| = 1$ and $\lambda > 1$, we thus end up with

$$\lambda = 1 + \sqrt{2}, \qquad \text{and} \qquad p_0 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Next step is therefore $x_1 = x_0 + p_0 = p_0$.

2 Let

$$f(x) = \tfrac{1}{2}x_1^2 + x_2^2,$$

put $x_0 = (1, 1)$, and define the model function $m(p) = f(x_0) + g^\top p + \frac{1}{2}p^\top Bp$ with $g = \nabla f(x_0)$ and $B = \nabla^2 f(x_0)$.

a) Compute explicitly the next step p in the trust region method using values of $\Delta = 2$ and $\Delta = 5/6$.

Solution: Note first that

$$g = \nabla f(x_0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{and} \qquad B = \nabla^2 f(x_0) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

and that the unconstrained minimizer of m equals $p_0^{\mathrm{B}} = -B^{-1}g = -(1,1)$. When $\Delta = 2$, this point is feasible—indeed, $\|p_0^{\mathrm{B}}\| = \sqrt{2} < 2$—and hence, we compute the next step with $p_0 = p_0^{\mathrm{B}}$ as $x_1 = x_0 + p_0 = (0,0)$, which turns out to be the global minimizer of f.

If, however, $\Delta = 5/6$, then (1) from Theorem 4.1 in N&W implies that

$$p_0 = -\begin{bmatrix} 1/(1+\lambda) \\ 2/(2+\lambda) \end{bmatrix}$$

for some $\lambda \geq 0$. We cannot have $\lambda = 0$, because then $p_0 = p_0^{\mathrm{B}}$, which is infeasible. Thus $\lambda > 0$ and $\|p_0\| = \Delta = 5/6$ by complementarity condition (2). Written out and simplifying, the latter equation becomes

$$\begin{aligned} 0 &= 25\lambda^4 + 150\lambda^3 + 145\lambda^2 - 132\lambda - 188 \\ &= (\lambda - 1)(25\lambda^3 + 175\lambda^2 + 300\lambda + 188). \end{aligned}$$

Since the second factor in the last expression is positive for all $\lambda \geq 0$, we infer that $\lambda = 1$ is the only possibility. This gives

$$p_0 = (-1/2, -2/3) \qquad \text{and} \qquad x_1 = x_0 + p_0 = (1/2, 1/3).$$

(Note that condition (3) is automatically satisfied because B is positive definite.)

b) Compute for all $\Delta > 0$ the next step in the dogleg method.

Solution: If $\Delta \geq 2$, the full step $p_0 = p_0^{\mathrm{B}}$ is feasible, yielding $x_1 = x_0 + p_0 = (0,0)$. Next, the steepest descent step equals

$$p_0^{\mathrm{U}} = -\frac{g^\top g}{g^\top B g} g = -\begin{bmatrix} 5/9 \\ 10/9 \end{bmatrix}$$

and satisfies $\|p_0^{\mathrm{U}}\| = 5\sqrt{5}/9 \approx 1.24$. If $\Delta \leq \|p_0^{\mathrm{U}}\|$, the dogleg method chooses p_0 to lie on the "steepest descent trajectory", scaled to lie on the boundary of the trust-region, so that

$$p_0 = \frac{\Delta}{\|p_0^{\mathrm{U}}\|} p_0^{\mathrm{U}} = -\frac{\Delta}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

This yields a new step $x_1 = \left(1 - \frac{\Delta}{\sqrt{5}}, 1 - \frac{2\Delta}{\sqrt{5}}\right)$. Observe that for $\Delta = 5/6$, this gives $x_1 \approx (0.63, 0.25)$, which is not too far from the optimal x_1 found in the previous problem.

For the remaining case $5\sqrt{5}/9 < \Delta < 2$, we follow the dogleg path

$$p(\tau) = p_0^{\mathrm{U}} + \tau \left(p_0^{\mathrm{B}} - p_0^{\mathrm{U}}\right), \qquad \tau \in (0,1)$$

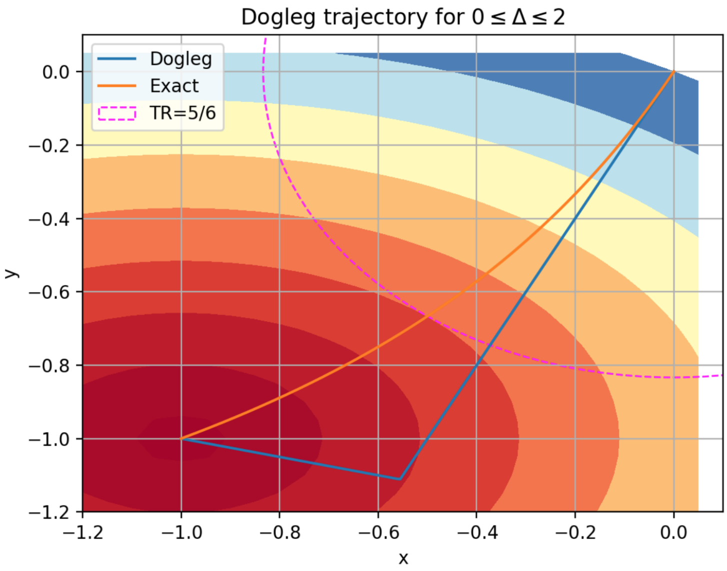

Figure 1: Comparion of the dogleg trajetory vs the exact solutions to the trust region problem.

until it hits the boundary of the trust-region, that is, when

$$\begin{aligned}\Delta^2 = \|p(\tau)\|^2 &= \|p_0^{\mathrm{U}}\|^2 + 2\tau\left(p_0^{\mathrm{B}} - p_0^{\mathrm{U}}\right)^\top p_0^{\mathrm{U}} + \tau^2\|p_0^{\mathrm{B}} - p_0^{\mathrm{U}}\|^2 \\ &= \frac{17}{81}\tau^2 + \frac{20}{81}\tau + \frac{125}{81}\end{aligned}$$

Solving this quadratic equation with respect to τ gives

$$\tau = -\frac{10}{17} + \frac{9}{17}\sqrt{17\Delta^2 - 25},$$

where the other solution has been discarded since it results in $\tau < 0$. Next step is therefore $x_1 = x_0 + p(\tau)$, with τ as above.

See Figure 1 which plots both the dogleg trajectory and the exact solution to the trust-region problem for $0 \leq \Delta \leq 2$. The plot has been obtained using mostly symbolic computations from `sympy`; however we also used numerical root finding to solve the Trust Region subproblem exactly, since this is faster. See the source code on the wiki page.

3 Find (and simplify, if possible) the dual of the linear programme

$$\min c^T x \qquad \text{subject to } Ax \geq b,\ x \geq 0.$$

Solution: Introducing the Lagrangian

$$\begin{aligned}\mathcal{L}(x, \lambda, s) &= c^\top x - \lambda^\top(Ax - b) - s^\top x \\ &= b^\top\lambda + (c - A^\top\lambda - s)^\top x,\end{aligned}$$

the dual problem is defined as

$$\max_{\lambda\geq 0, s\geq 0} \min_x \mathcal{L}(x,\lambda,s). \tag{$\star$}$$

Since

$$\min_x \mathcal{L}(x,\lambda,s) = \begin{cases} -\infty & \text{if } A^\top\lambda + s \neq c; \\ b^\top\lambda & \text{if } A^\top\lambda + s = c, \end{cases}$$

we see that ($\star$) is equivalent to the problem

$$\max_{\lambda\geq 0, s\geq 0} b^\top\lambda \qquad \text{subject to} \qquad A^\top\lambda + s = c.$$

Interpreting s as a slack variable, we can further simplify the dual problem to

$$\max_\lambda b^\top\lambda \qquad \text{subject to} \qquad A^\top\lambda \leq c,\ \lambda\geq 0.$$

4 Find the dual of the linear optimisation problem

$$5x_1 + 3x_2 + 4x_3 \to \min \qquad \text{subject to} \begin{cases} x_1 + x_2 + x_3 = 1, \\ \qquad\qquad x_i \geq 0, \quad i = 1,2,3, \end{cases}$$

and compute its (i.e., the *dual's*) solution.

Solution: Abstractly, the linear optimisation problem may be written as

$$\min_x c^\top x \qquad \text{subject to} \qquad Ax = b,\ x\geq 0,$$

and its dual is, similarly as in the previous exercise,

$$\max_\lambda b^\top\lambda \qquad \text{subject to} \qquad A^\top\lambda \leq c.$$

Note that the constraint $\lambda \geq 0$ is not present in this case (why?). With

$$c = (5,3,4), \quad A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad b = 1,$$

the actual dual becomes

$$\max \lambda \qquad \text{subject to} \qquad \lambda\leq 5,\ \lambda\leq 3,\ \lambda\leq 4,$$

with obvious solution $\lambda^* = 3$. Observe also that this gives us a convenient way of solving the primal problem: since only the second constraint is active, we have $x_1 = 0 = x_3$, and so the equality constraint of the primal problem yields that $x_2 = 1$.