Exercises (10 tasks, 5 points in total)

Each correctly solved task gives 0.5 points.

1. Let
$$f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2.$$
At $x = (0, -1)$ draw the contour lines of the quadratic model
$$m_k(p) = f_k + g_k^T p + \tfrac{1}{2} p^T B_k p,$$
where B_k is the Hessian of f_k, $g_k = \nabla f(x_k)$ and $f_k = f(x_k)$. Draw the family of solutions of the trust-region subproblem
$$\min_{p \in \mathbb{R}^n} \; m_k(p) \quad \text{s.t. } \|p\| \leq \Delta_k,$$
as the trust region radius Δ varies from 0 to 2. Repeat this at $x = (0, 0.5)$.

2. Write a program that implements the dogleg method. Choose B_k to be the exact Hessian. Apply it to solve Rosenbrock's function
$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$
Experiment with the update rule for the trust region by changing the constants in the trust-region algorithm below, or by designing your own rules.

   Algorithm (Trust Region):

   (a) Given $\hat{\Delta} > 0$, $\Delta_0 \in (0, \hat{\Delta})$, and $\eta \in [0, \frac{1}{4})$.

   (b) For $k = 0, 1, 2, \ldots$:

      i. Obtain p_k by (approximately) solving the subproblem above.

      ii. Evaluate ρ_k.

      iii. If $\rho_k < \frac{1}{4}$, set $\Delta_{k+1} = \frac{1}{4}\Delta_k$.

      iv. Else if $\rho_k > \frac{3}{4}$ and $\|p_k\| = \Delta_k$, set $\Delta_{k+1} = \min(2\Delta_k, \hat{\Delta})$.

      v. Else set $\Delta_{k+1} = \Delta_k$.

      vi. If $\rho_k > \eta$, set $x_{k+1} = x_k + p_k$, else $x_{k+1} = x_k$.

3. Theorem 14 (see Lecture Note 4) states that the sequence $\{\|g_k\|\}$ has an accumulation point at zero. Show that if the iterates x_k stay in a bounded set B, then there is a limit point x_∞ of the sequence $\{x_k\}$ such that

$$g(x_\infty) = 0.$$

4. Show that τ_k defined by

$$\tau_k = \begin{cases} 1, & \text{if } g_k^T B_k g_k \le 0, \\ \min\left(\dfrac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k}, 1 \right), & \text{otherwise,} \end{cases}$$

does indeed identify the minimizer of m_k along the direction $-g_k$.

5. The Cauchy–Schwarz inequality states that for any vectors u and v,

$$|u^T v|^2 \le (u^T u)(v^T v),$$

with equality only when u and v are parallel. When B is positive definite, use this inequality to show that

$$\gamma \stackrel{\text{def}}{=} \frac{\|g\|^4}{(g^T B g)(g^T B^{-1} g)} \le 1,$$

with equality only if g, Bg, and $B^{-1}g$ are parallel.

6. Show that the following two root-finding updates are equivalent:

$$\lambda^{(\ell+1)} = \lambda^{(\ell)} - \frac{\phi_2(\lambda^{(\ell)})}{\phi_2'(\lambda^{(\ell)})}, \tag{A}$$

and

$$\lambda^{(\ell+1)} = \lambda^{(\ell)} + \left(\frac{\|p_\ell\|}{\|q_\ell\|} \right)^2 (\|p_\ell\| - \Delta), \tag{B}$$

using the identities

$$\frac{d}{d\lambda}\left(\|p(\lambda)\|^{-1} \right) = -\tfrac{1}{2} \|p(\lambda)\|^{-3} \frac{d}{d\lambda} \|p(\lambda)\|^2,$$

$$\frac{d}{d\lambda} \|p(\lambda)\|^2 = -2 \sum_{j=1}^{n} \frac{(q_j^T g)^2}{(\lambda_j + \lambda)^3},$$

and

$$\|q\|^2 = \sum_{j=1}^{n} \frac{(q_j^T g)^2}{(\lambda_j + \lambda)^3}.$$

7. Derive the solution of the two-dimensional subspace minimization problem in the case where B is positive definite.

8. Show that if B is any symmetric matrix, then there exists $\lambda \geq 0$ such that $B + \lambda I$ is positive definite.

9. Verify that the definitions

$$p_k^S = -\frac{\Delta_k}{\|D^{-1} g_k\|} D^{-2} g_k,$$

and

$$\tau_k = \begin{cases} 1, & \text{if } g_k^T D^{-2} B_k D^{-2} g_k \leq 0, \\ \min\left(\dfrac{\|D^{-1} g_k\|^3}{\Delta_k g_k^T D^{-2} B_k D^{-2} g_k}, 1 \right), & \text{otherwise,} \end{cases}$$

are valid for the Cauchy point in the case of an elliptical trust region, where D is a diagonal scaling matrix with positive diagonal elements.

10. Consider the trust-region subproblem

$$\min_{p \in \mathbb{R}^n} \; m_k(p) = f_k + g_k^T p + \tfrac{1}{2} p^T B_k p \quad \text{s.t. } \|p\| \leq \Delta_k.$$

Suppose B_k is positive definite. Prove that if the unconstrained minimizer $p_B = -B_k^{-1} g_k$ satisfies $\|p_B\| \leq \Delta_k$, then p_B is also the solution of the trust-region subproblem. Otherwise, the solution lies on the boundary $\|p\| = \Delta_k$.