Answer Sheet

1. Compute the gradient and Hessian of the function $q(x) = \frac{1}{2}x^T Ax + b^T x$, where A is symmetric.

   Sol. Gradient: $g(x) = Ax + b$; Hessian: $H(x) = A$.

2. Compute the gradient and Hessian of

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

   Show that $x^* = (1, 1)^T$ is the only local minimizer of this function, and that the Hessian at this point is positive definite. Write a program on trust region method with subproblems solved by the Dogleg method. Apply it to minimize this function. Choose $B^k = \nabla^2 f(x^k)$. Experiment with the update rule of trust region. Give the first two iterates.

   Pf. The gradient $g(x) = 0$ gives the unique solution $x^* = (1, 1)^T$, which thus is the only local minimizer of f. And the Hessian at x^* is $H(x^*) = [802, -400; -400, 200]$ which is positive definite. □

3. Apply SD method with exact line search to the problem:

$$\min_{x=(x_1,x_2)^T} \quad f(x) = 5x_1^2 + \frac{1}{2}x_2^2.$$

   Carry out two iterations, starting from $x^0 = (0.1, 1)^T$. Think about how $\{x^k\}$ will behave.

   Sol. The gradient is $g = (10x_1, x_2)^T$, and the Hessian is $H = \mathrm{diag}(10, 1)$. Use the step size $\alpha_k = \frac{(g^k)^T g^k}{(g^k)^T H g^k}$ for the steepest descent method. Here is the table which records the iteration information.

k	$(x^k)^T$	$(g^k)^T$	α_k
0	$(\frac{1}{10}, 1)^T$	$(1, 1)^T$	$\frac{2}{11}$
1	$\frac{9}{11}(-\frac{1}{10}, 1)^T$	$\frac{9}{11}(-1, 1)^T$	$\frac{2}{11}$
2	$\frac{81}{11}(\frac{1}{10}, 1)^T$	$\frac{81}{121}(1, 1)^T$	$\frac{2}{11}$

   Conjecture:

$$x^k = \left(\frac{9}{11}\right)^k \left(\frac{(-1)^k}{10}, 1\right)^T.$$

   We can prove this conjecture by the way of induction.

4. Assume that there exsits $M > 0$ such that $\|B_k\|\|B_k^{-1}\| \leq M$ where B_k is symmetric and positive definite and $\|\cdot\|$ refers to the Euclidean 2-norm. Show that if $d_k = -B_k^{-1}\nabla f_k$,

$$\frac{-d_k^T \nabla f_k}{\|d_k\|\|\nabla f_k\|} \geq \frac{1}{M}.$$

Pf. It is easy to have

$$\begin{aligned}\frac{-d_k^{\nabla} f_k}{\|d_k\|\|\nabla f_k\|} =&\frac{d_k^T B_k d_k}{\|d_k\|\|B_k d_k\|}\\ \geq&\frac{\|B_k^{\frac{1}{2}} d_k\|^2}{\|B_k^{-1}\|\|B_k d_k\|^2}\\ \geq&\frac{\|B_k^{\frac{1}{2}} d_k\|^2}{\|B_k^{-1}\|\|B_k\|\|B_k^{\frac{1}{2}} d_k\|^2}\\ \geq&\frac{1}{M},\end{aligned}$$

where the third inequality follows from

$$\|B_k d_k\|^2 = d_k^T B_k^2 d_k = (B_k^{\frac{1}{2}} d_k)^T B_k (B_k^{\frac{1}{2}} d_k) \leq \|B_k\|\|B_k^{\frac{1}{2}} d_k\|^2.$$

5. Apply CG method with exact line search to solve

$$\min_{x\in\mathbb{R}^2} \quad \frac{1}{2}x^T Ax + b^T x,$$

starting from $x_0 = (2,1)^T$. Here $A = [4,1;1,3]$ and $b = -(1,2)^T$.
Sol.

(i) $k = 0$.

$x_0 = (2,1)^T$, $g_0 = Ax_0 + b = (8,3)^T$, $d_0 = -g_0$, $\alpha_0 = -\frac{g_0^T d_0}{d_0^T A d_0} = \frac{73}{331}$.

(ii) $k = 1$.

$x_1 = x_0 + \alpha_0 d_0 = [0.2356, 0.3384]^T$, $g_1 = Ax_1 + b = [0.2810, -0.7492]^T$,
$\beta_0 = \frac{g_1^T A d_0}{d_0^T A d_0} = 0.0088$,
$d_1 = -g_1 + \beta_0 d_0 = (-0.3511, 0.7229)^T$, $\alpha_1 = -\frac{g_1^T d_1}{d_1^T A d_1} = 0.4122$,
$x_2 = x_1 + \alpha_1 d_1 = (0.0909, 0.6364)^T$.

If exact arithmetic is to be used in this example instead of limited-precision, the exact solution would theoretically have been reached after $n = 2$ iterations.