Please identify anyone, whether or not they are in the class, with whom you discussed your homework. This problem is worth 1 point, but on a multiplicative scale (multiplied by the rest). Collaboration is encouraged.

For all problems in this homework set, $\boldsymbol{Q} \in \mathbb{R}^{n\times n}_{\text{sym}}$ is set to be symmetric positive definite, i.e., $\boldsymbol{v}^\top \boldsymbol{Q}\boldsymbol{v} > 0$ for $\boldsymbol{v} \neq \boldsymbol{0}$, and $\boldsymbol{Q}^\top = \boldsymbol{Q}$. Its eigenvalues are $\lambda_n \geq \cdots \geq \lambda_1 > 0$ such that $\lambda_n = \lambda_{\max}(\boldsymbol{Q})$ and $\lambda_1 = \lambda_{\min}(\boldsymbol{Q}) > 0$. The objective is to use an iterative method (steepest descent or CG) $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$ with an initial guess of $\boldsymbol{x}^{(0)}$ to find the minimizer $\boldsymbol{x}^* = \boldsymbol{Q}^{-1}\boldsymbol{b}$ of

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) := \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} \right\}.$$

1. (5 points) (Lottery ticket) Show that if the initial guess $\boldsymbol{x}^{(0)}$ is such that $\boldsymbol{x}^{(0)} - \boldsymbol{x}^*$ is parallel to one of the eigenvectors of Q (i.e., there exists an eigenvector $\boldsymbol{v}_i$ such that $\boldsymbol{x}^{(0)} - \boldsymbol{x}^* = \beta \boldsymbol{v}_i$ for some $\beta \neq 0$), then the steepest descent method (linear search) will reach to $\boldsymbol{x}^*$ in one iteration.

2. (5 points) Consider a set of nonzero vector $\{\boldsymbol{p}_0, \dots, \boldsymbol{p}_n\}$ such that for $0 \leq i, j \leq n$

   $$(\boldsymbol{p}_i, \boldsymbol{p}_j)_{\boldsymbol{Q}} := (\boldsymbol{Q}\boldsymbol{p}_i, \boldsymbol{p}_j) = \boldsymbol{p}_i^\top \boldsymbol{Q}\boldsymbol{p}_j = 0 \quad \text{if } \ i \neq j,$$

   then we say this $\boldsymbol{Q}$-orthogonal set of vector is *conjugate* with respect to $\boldsymbol{Q}$, or $\boldsymbol{Q}$-conjugate. Show that these vectors $\{\boldsymbol{p}_0, \dots, \boldsymbol{p}_n\}$ are linearly independent, i.e., they automatically form a basis for their span, and this implies that an $n \times n$ matrix has most n conjugate directions. (Hint: use proof by contradiction, $\|\cdot\|_{\boldsymbol{Q}} := \sqrt{(\cdot,\cdot)_{\boldsymbol{Q}}}$ being a norm, and a few linear algebra tricks)

3. (10 points) Consider $\boldsymbol{x} = (x_1, x_2)^\top$ and the following quadratic function $f : \mathbb{R}^2 \to \mathbb{R}$

   $$f(x_1, x_2) = \frac{5}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1x_2 - 3x_1 - x_2$$

   (a) Express f in the form of $\frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x}$.

   (b) Find the minimizer of f using the conjugate gradient algorithm by hand using the initial guess $x^{(0)} = (0, 0)^\top$.

   (c) Calculate the minimizer of f analytically from $\boldsymbol{Q}$ and $\boldsymbol{b}$, and check it with your answer in part (b).

4. (5 points) (Generalized CG) Let $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, and denote as usual $\boldsymbol{g}^{(k)} := \nabla f(\boldsymbol{x}^{(k)})$, and suppose the search directions are generated according to

   $$\boldsymbol{d}^{(k+1)} = \alpha_k \boldsymbol{g}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)}, \quad k = -1, 0, 1, 2, \dots$$

   where α_k, β_k are real constants and by convention let $\boldsymbol{d}^{(-1)} = \boldsymbol{0}$. Note the standard CG corrects the negative gradient and $\alpha_k = -1$ for all k. Let the Krylov subspace of order $(k+1)$ be $\mathcal{V}_{k+1} := \operatorname{span}\{\boldsymbol{b}, \boldsymbol{Q}\boldsymbol{b}, \dots, \boldsymbol{Q}^k\boldsymbol{b}\}$. Show that $\boldsymbol{d}^{(k)} \in \mathcal{V}_{k+1}$ and $\boldsymbol{x}^{(k)} \in \mathcal{V}_k$. (Hint: use proof by induction, note $\mathcal{V}_0 = \{\boldsymbol{0}\}$ and $\mathcal{V}_1 = \operatorname{span}\{\boldsymbol{b}\}$)