Math 408 Homework Set 3 Solutions

This homework set will focus on the optimization problem

$$\mathcal{Q} \qquad \min_{x\in\mathbb{R}^n} \frac{1}{2}x^T Hx + g^T x \ ,$$

where $H \in \mathbb{R}^{n\times n}$ is symmetric and $g \in \mathbb{R}^n$.

(1) Each of the following functions can be written in the form $f(x) = \frac{1}{2}x^T Hx + g^T x$ with H symmetric. For each of these functions what are H and g

(a) $f(x) = x_1^2 - 4x_1 + 2x_2^2 + 7$

Solution

$$H = 2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad g = (-4,\ 0)^T$$

(c) $f(x) = x_1^2 - 2x_1x_2 + \frac{1}{2}x_2^2 - 8x_2$

Solution

$$H = 2\begin{bmatrix} 1 & -1 \\ -1 & 1/2 \end{bmatrix} \quad \text{and} \quad g = (0,\ -8)^T$$

(d) $f(x) = (2x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$

Solution:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad H = 2A^TA = 2\begin{bmatrix} 4 & -2 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad \text{and} \quad g = -A^Tb = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

(e) $f(x) = x_1^2 + 16x_1x_2 + 4x_2x_3 + x_2^2$

Solution:

$$H = 2\begin{bmatrix} 1 & 8 & 0 \\ 8 & 1 & 2 \\ 0 & 2 & 0 \end{bmatrix} \quad \text{and} \quad g = 0$$

(2) Let $H \in \mathcal{S}^n_{++}$ and g, $\bar{x}$, $d \in \mathbb{R}^n$ with $d \neq 0$, and consider the functions $f : \mathbb{R}^n \to \mathbb{R}$ and $\phi : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) := \frac{1}{2}x^T Hx + g^T x \ \text{ and } \ \phi(t) := f(\bar{x} + td).$$

What is the solution to the problem $\min_{t\in\mathbb{R}} \phi(t)$?

Solution:

$$\phi(t) = f(\bar{x}) + t(H\bar{x} + g)^T d + \frac{t^2}{2}d^T Hd \ \text{ and } \ \phi'(t) = td^T Hd + (H\bar{x} + g)^T d.$$

Since ϕ is a quadratic function of t with $\phi''(t) = d^T Hd > 0$, the global minimizer of ϕ occurs when $\phi'(t) = 0$, that is, when $t = -\frac{(H\bar{x}+g)^T d}{d^T Hd}$.

(3) Let $\alpha,\ \beta \in \mathbb{R}$ and define $\phi : \mathbb{R} \to \mathbb{R}$ by $\phi(t : \alpha, \beta) := \frac{\alpha}{2}t^2 + \beta t$. Compute the values $\bar\phi(\alpha,\beta) := \min\{\phi(t : \alpha,\beta)\ |0 \le t \le 1\}$ and the set $\Gamma(\alpha,\beta) := \operatorname{argmin}\{\phi(t;\alpha,\beta)\ |0 \le t \le 1\}$ for all possible values of $(\alpha,\beta) \in \mathbb{R}^2$.

Solution:

$$\Gamma(\alpha,\beta) := \begin{cases} \{0\} & ,\ \alpha = 0,\ \beta > 0, \\ [0,1] & ,\ \alpha = 0,\ \beta = 0, \\ \{1\} & ,\ \alpha = 0,\ \beta < 0, \\ \{0\} & ,\ \alpha < 0,\ \frac{-\alpha}{2} < \beta, \\ \{0,1\} & ,\ \alpha < 0,\ \frac{-\alpha}{2} = \beta, \\ \{1\} & ,\ \alpha < 0,\ \beta < \frac{-\alpha}{2}, \\ \{0\} & ,\ \alpha > 0,\ 0 \le \beta, \\ \{\frac{-\beta}{\alpha}\} & ,\ \alpha > 0,\ 0 < -\beta < \alpha, \\ \{1\} & ,\ \alpha > 0,\ \alpha \le -\beta. \end{cases} \qquad \bar\phi(\alpha,\beta) = \begin{cases} 0 & ,\ \alpha = 0,\ 0 \le \beta, \\ \beta & ,\ \alpha = 0,\ \beta < 0, \\ 0 & ,\ \alpha < 0,\ \frac{-\alpha}{2} \le \beta, \\ \frac{\alpha}{2} + \beta & ,\ \alpha < 0,\ \beta < \frac{-\alpha}{2}, \\ 0 & ,\ \alpha > 0,\ 0 \le \beta, \\ -\frac{\beta^2}{2\alpha} & ,\ \alpha > 0,\ 0 < -\beta < \alpha, \\ \frac{\alpha}{2} + \beta & ,\ \alpha > 0,\ \alpha, -\beta. \end{cases}$$

(4) For each of the matrices H and vectors g below determine the optimal value in $\mathcal{Q}$. If an optimal solution exists, compute the complete set of optimal solutions.

Solution: For this analysis our primary tool is Theorem 1.1 on page 26 of the course text.

(a)

$$H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

Solution: The principal minors of H are 2, 5, and 8 which are all positive. Hence, by Proposition 3.1 on page 30 of the course notes, H is positive definite. Therefore the unique optimal solution is given by $-H^{-1}g = (-2, 1, -1)^T$.

(b)

$$H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

Solution: We have $e_3^T H e_3 = -2$. Set $f(x) := \frac{1}{2}x^T H x + g^T x$. Then $f(te_3) = -t^2 + t$ so that $f(te_3) \downarrow -\infty$ as $t \uparrow +\infty$, so the optimal value is $-\infty$.

(c)

$$H = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}.$$

Solution: Observe that the principal minors of H are 5, 1, and 0. So H is not positive definite. We also know that the set of solutions to the equation $0 = Hx + g$ is

$$x = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -3 \\ -1 \end{pmatrix} \quad \forall\ t \in \mathbb{R}.$$

Also observe that $\langle g\,,\,(1,\ -3,\ -1)\rangle = 0$ If we can show that H is positive semidefinite, then this is the set of optimal solutions to $\mathcal{Q}$. For this we use part (1) of Lemma 3.1 on page 29 of the course text. First observe that H is of the form $\begin{bmatrix} \hat{H} & u \\ u^t & \alpha \end{bmatrix}$, where $\hat{H} := \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$, $u = (-1,\ -1)^T$ and $\alpha = 2$. Our computation of the principal minors shows that the matrix $\hat{H}$ is positive definite. Solving the equation $\hat{H}z = u$ gives $z = (1,\ -3)^T$, and $z^T\hat{H}z = 2$. Consequently, H is positive definite, and we have completely specified all optimal solutions to $\mathcal{Q}$.

(5) Consider the matrix $H \in \mathbb{R}^{3\times 3}$ and vector $g \in \mathbb{R}^3$ given by

$$H = \begin{bmatrix} 1 & 4 & 1 \\ 4 & 20 & 2 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} .$$

Does there exists a vector $u \in \mathbb{R}^3$ such that $f(tu) \xrightarrow{t\uparrow\infty} -\infty$? If yes, construct u.

Solution: We can show that H is positive definite using the technique described in part (c) of problem 3 above ($\hat{H} = \begin{bmatrix} 1 & 4 \\ 4 & 20 \end{bmatrix}$ is positive definite, solve $\hat{H}z = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ to get $z = (3,\ -\frac{1}{2})$ and note that $z^T\hat{H}z = 2$). However, the system $0 = Hx + g$ is easily seen to be inconsistent, so the optimal value is $-\infty$. To obtain the desired vector u, we need to obtain a vector in the null space of H whose inner product with g is non-zero. We compute the null space of H to be the span of the vector $u := (-6,\ 1,\ 2)^T$. The inner product of g with u is -6. Hence, if $f(x) := \frac{1}{2}x^THx + g^Tx$, then $f(tu) = -6t$ which goes to $-\infty$ as $t \uparrow +\infty$.

(6) A mapping $\langle \cdot\,,\,\cdot \rangle : \mathbb{R}^n \to \mathbb{R}^n$ is said to be an inner product on $\mathbb{R}^n$ is for all $x, y, z \in \mathbb{R}^n$

(i)	$\langle x\,,\,x\rangle \geq 0$	Non-Negative
(ii)	$\langle x\,,\,x\rangle = 0 \Leftrightarrow x = 0$	Positive
(iii)	$\langle x+y\,,\,z\rangle = \langle x\,,\,z\rangle + \langle y\,,\,z\rangle$	Additive
(iv)	$\langle \alpha x\,,\,y\rangle = \alpha\langle x\,,\,y\rangle\ \forall\, \alpha \in \mathbb{R}$	Homogeneous
(v)	$\langle x\,,\,y\rangle = \langle y\,,\,x\rangle$	Symmetric

Two vectors $x, y \in \mathbb{R}^n$ are said to be orthogonal in the inner product $\langle \cdot\,,\,\cdot \rangle$ if $\langle x\,,\,y\rangle = 0$

Unless otherwise specified, we use the notation $\langle x\,,\,y\rangle$ to designate the usual Euclidean inner product:

$$\langle x\,,\,y\rangle = \sum_{i=1}^{n} x_i y_i \ .$$

(a) Let $\langle x\,,\,y\rangle$ be the Euclidean inner product on $\mathbb{R}^n$. Given $A \in \mathbb{R}^{n\times n}$, show that $A = 0$ if and only if

$$\langle x\,,\,Ay\rangle = 0 \quad \forall\, x, y \in \mathbb{R}^n \ .$$

Solution: If $A = 0$, then $\langle x\,,\,Ay\rangle = 0$. Conversely, note that $A_{ij} = \langle e_i\,,\,Ae_j\rangle = 0$, where e_i is the vector with all zeros and a 1 in the ith position, so all the entries of A must be zero.

(b) Let $H \in \mathbb{R}^{n\times n}$ be symmetric and positive definite (i.e. $H = H^T$ and $x^THx > 0\ \forall\, x \in \mathbb{R}^n \setminus \{0\}$). Show that the bi-linear form given by

$$\langle x\,,\,y\rangle_H = x^THy \quad \forall\, x, y \in \mathbb{R}^n$$

defines an inner product on $\mathbb{R}^n$.

Solution: Properties (i) and (ii) are immediate from the definition of positive definite matrices. Properties (iii) and (iv) follow from the linearity of matrix multiplication. Property (v) follows from the symmetry of H.

(c) Every inner product defines a transformation on the space of linear operators called the *adjoint*. For the Euclidean inner product on $\mathbb{R}^n$, this is just the usual transpose. Given a linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$, the adjoint is defined by the relation

$$\langle y\,,\, Mx\rangle = \langle M^* y\,,\, x\rangle, \quad \text{for all } x, y, \in \mathbb{R}^n.$$

The inner product given above, $\langle \cdot\,,\, \cdot\rangle_H$, also defines an adjoint mapping which we can denote by M^{T_H}. Show that

$$M^{T_H} = H^{-1} M^T H\ .$$

Solution:

$\langle y\,,\, Mx\rangle_H = \langle y\,,\, HMx\rangle = \ \langle M^T Hy\,,\, x\rangle = \langle M^T Hy\,,\, H^{-1}x\rangle_H = \langle H^{-1}M^T Hy\,,\, x\rangle_H$

Comparing the leftmost and rightmost expressions gives the result.

(d) The matrix $P \in \mathbb{R}^{n\times n}$ is said to be a projection if $P^2 = P$. Clearly, if P is a projection, then so is $I - P$. The subspace $P\mathbb{R}^n = \mathrm{Ran}\,(P)$ is called the subspace that P projects onto. A projection is said to be orthogonal with respect to a given inner product $\langle \cdot\,,\, \cdot\rangle$ on $\mathbb{R}^n$ if and only if

$$\langle (I-P)x\,,\, Py\rangle = 0 \quad \forall\ x, y \in \mathbb{R}^n\ ,$$

that is, the subspaces $\mathrm{Ran}\,(P)$ and $\mathrm{Ran}\,(I-P)$ are orthogonal in the inner product $\langle \cdot\,,\, \cdot\rangle$. Show that the projection P is orthogonal with respect to the inner product $\langle \cdot\,,\, \cdot\rangle_H$ (defined above), where $H \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, if and only if

$$P = H^{-1}P^T H\ .$$

Solution: Note that

$$\langle (I-P)x\,,\, Py\rangle_H = \langle (I-P)x\,,\, HPy\rangle = \langle P^T H(I-P)x\,,\, y\rangle$$

If $P = H^{-1}P^T H$, then $P^T H(I-P) = HPH^{-1}H(I-P) = HP(I-P) = H(P-P) = 0$, so the result follows.

For the converse, we have $\langle P^T H(I-P)x\,,\, y\rangle = 0$, so $P^T H(I-P) = 0$ by (a). Then $P^T H = P^T HP$, so $P^T = P^T HPH^{-1}$. Taking the transpose, we have $P = H^{-1}(P^T HP) = H^{-1}P^T H$ since $P^T H = P^T HP$.