AMATH/MATH 516
FIRST HOMEWORK SET SOLUTIONS

The purpose of this problem set is to have you brush up and further develop your multi-variable calculus and linear algebra skills. The problem set will be very difficult for some and straightforward for others. If you are having any difficulty, please feel free to discuss the problems with me at any time. Don't delay in starting work on these problems!

1. Let Q be an $n \times n$ symmetric positive definite matrix. The following fact for symmetric matrices can be used to answer the questions in this problem.

   Fact: *If M is a real symmetric $n \times n$ matrix, then there is a real orthogonal $n \times n$ matrix U ($U^T U = I$) and a real diagonal matrix $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ such that $M = U\Lambda U^T$.*

   (a) Show that the eigenvalues of Q^2 are the square of the eigenvalues of Q.

   Note that λ is an eigenvalue of Q if and only if there is some vector v such that $Qv = \lambda v$. Then $Q^2 v = Q\lambda v = \lambda^2 v$, so λ^2 is an eigenvalue of Q^2.

   (b) If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigen values of Q, show that

   $$\lambda_n \|u\|_2^2 \leq u^T Q u \leq \lambda_1 \|u\|_2^2 \ \forall \ u \in I\!R^n.$$

   We have $u^T Q u = u^T U \Lambda U u = \sum_{i=1}^n \lambda_i \|u\|_2^2$. The result follows immediately from bounds on the λ_i.

   (c) If $0 < \underline{\lambda} < \bar{\lambda}$ are such that

   $$\underline{\lambda}\|u\|_2^2 \leq u^T Q u \leq \bar{\lambda}\|u\|_2^2 \ \forall \ u \in I\!R^n,$$

   then all of the eigenvalues of Q must lie in the interval $[\underline{\lambda}, \bar{\lambda}]$.

   Take u_1 to be the eigenvector for λ_1. Then $u_1^T Q u_1 = \lambda_1 \|u_1\|_2^2$, so $\lambda_1 \leq \bar{\lambda}$. The analogous argument applied to u_n, the eigenvector associated to λ_n, gives $\underline{\lambda} \leq \lambda_n$.

   (d) Let $\underline{\lambda}$ and $\bar{\lambda}$ be as in Part (c) above. Show that

   $$\underline{\lambda}\|u\|_2 \leq \|Qu\|_2 \leq \bar{\lambda}\|u\|_2 \ \forall \ u \in I\!R^n \ .$$

   Hint: $\|Qu\|_2^2 = u^T Q^2 u$.

   From part (a), we have that the eigenvalues of Q^2 are λ_i^2, where λ_i are the eigenvalues of Q. From part b we have that $\lambda_n^2 \|u\|_1^2 \leq \|Qu\|_2^2 \leq \lambda_1^2 \|u\|_2^2 \ \forall u \in \mathbf{R}^n$, and the result follows immediately.

2. Consider the quadratic function $f : I\!R^n \mapsto I\!R$ given by

   $$f(x) := \frac{1}{2} x^T Q x - a^T x + \alpha \ ,$$

   where $Q \in I\!R^{n \times n}$, $a \in I\!R^n$, and $\alpha \in I\!R$.

   (a) Write expressions for both $\nabla f(x)$ and $\nabla^2 f(x)$. Since it is not assumed that f is symmetric, be careful in how you express $\nabla^2 f(x)$.

   $\nabla f(x) = \frac{1}{2}(Qx + Q^T x) - a$, *and* $\nabla^2 f(x) = \frac{1}{2}(Q + Q^T)$.

(b) If it is further assumed that Q is symmetric, what is $\nabla^2 f$?

It is Q.

(c) State first– and second–order necessary conditions for optimality in the problem $\min\{f(x) : x \in I\!R^n\}$.

We want $\nabla f(x) = 0$ and Q to be positive semidefinite.

(d) State a sufficient condition on the matrix Q under which the problem $\min f$ has a unique global solution and then display this solution in terms of the data Q and a.

Q should be positive definite; in which case the solution is given by $\frac{1}{2}(Q + Q^T)^{-1}a$. We might as well assume Q is symmetric.

3. Consider the linear equation

$$Ax = b,$$

where $A \in I\!R^{m \times n}$ and $b \in I\!R^m$. When $n < m$ it is often the case that this equation is over-determined in the sense that no solution x exists. In such cases one often attempts to locate a 'best' solution in a least squares sense. That is one solves the *linear least squares problem*

$$(\text{lls}) : \text{minimize } \frac{1}{2}\|Ax - b\|_2^2$$

for x. Define $f : I\!R^n \mapsto I\!R$ by

$$f(x) := \frac{1}{2}\|Ax - b\|_2^2.$$

(a) Show that f can be written as a quadratic function, that is, it can be written in the same form as the function of the preceding exercise.

$$\begin{aligned} f(x) &= \tfrac{1}{2}\langle Ax - b, Ax - b\rangle \\ &= \tfrac{1}{2}x^T A^T Ax - b^T Ax + \tfrac{1}{2}b^T b \end{aligned}$$

(b) What are $\nabla f(x)$ and $\nabla^2 f(x)$?

$$\nabla f(x) = A^T Ax - A^T b;,\ \nabla^2 f(x) = A^T A.$$

(c) Show that $\nabla^2 f(x)$ is positive semi-definite.

$$u^T A^T Au = \|Au\|_2^2 \geq 0.$$

(d) Show that $\mathrm{Nul}(A^T A) = \mathrm{Nul}(A)$ and $\mathrm{Ran}(AA^T) = \mathrm{Ran}(A)$.

It is clear that $Nul(A) \subset Nul(A^T A)$. For the other direction, if $w \in Nul(A^T A)$, then $A^T Ax = 0$ so that $0 = x^T A^T Ax = \|Ax\|^2$, or equivalently, $Ax = 0$.

By the Fundamental Theorem of the Alternative for matrices and the fact that $Nul(AA^T) = Nul(A^T)$ (obtained by replacing A with A^T in what was just proved), we have

$$Ran(AA^T) = Nul(AA^T)^{\perp} = Nul(A^T)^{\perp} = Ran(A).$$

(e) Show that a solution to (lls) must always exist.

The necessary and sufficient conditions for optimality are that $\nabla f = 0$ *and* $\nabla^2 f$ *is positive semidefinite (which we showed in part c) since this implies that* f *is convex. Now note that the equation* $A^T Ax = A^T b$ *always has a solution by part (d), since the ranges of the matrices* $A^T A$ *and* A^T *agree.*

(f) Provide a necessary and sufficient condition on the matrix A under which (lls) has a unique solution and then display this solution in terms of the data A and b.

By 2d), we know we need $\nabla^2 f$ *to be positive definite. Since we know* $\nabla^2 f = A^T A$ *is already positive semidefinite by 3c), we just need it to have a trivial nullspace. By 2d), we know* $A^T A$ *will have a trivial nullspace if* A *does. From this it immediately follows that* $Nul(A) = \{0\}$*, and so* $m \geq n$*. In this case, the solution is given by* $x = (A^T A)^{-1} A^T b$*.*

4. A mapping $\langle \cdot, \cdot \rangle : I\!R^n \mapsto I\!R^n$ is said to be an inner product on $I\!R^n$ is for all $x, y, z \in I\!R^n$

(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)	$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle \;\; \forall\, \alpha \in I\!R$	Homogeneous
(v)	$\langle x, y \rangle = \langle y, x \rangle$	Symmetric

Two vectors $x, y \in I\!R^n$ 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 $I\!R^n$. Given $A \in I\!R^{n \times n}$, show that $A = 0$ if and only if

$$\langle x, Ay \rangle = 0 \quad \forall\, x, y \in I\!R^n \ .$$

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* i*th position, so all the entries of* A *must be zero.*

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

$$\langle x, y \rangle_H = x^T H y \quad \forall\, x, y \in I\!R^n$$

defines an inner product on $I\!R^n$.

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 $I\!R^n$, this is just the usual transpose. Given a linear transformation $M : I\!R^n \mapsto I\!R^n$, the adjoint is defined by the relation

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

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

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

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

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

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

Note that

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

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 $\left\langle P^T H(I-P)x, y \right\rangle = 0$, 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 $P^T H = P^T HP$.

5. Consider the minimization problem

$$\mathcal{P}: \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \ , \end{array}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is assumed to be twice continuously differentiable, $A \in \mathbb{R}^{m\times n}$ has full rank with $m \le n$, and $b \in \mathbb{R}^m$. Set

$$P := I - A^T(AA^T)^{-1}A \ .$$

(a) Show that P is well–defined, that is, show that the matrix AA^T is non–singular.

From problem 4, we know that the nullspace of AA^T is the same as the nullspace of A^T. Since we are given that the matrix has full rank, it must have rank m, and since A^T maps from $\mathbf{R}^m \to \mathbf{R}^n$, it must have nullspace $\{0\}$. Then the square matrix AA^T is nonsingular.

(b) Show that P is the *orthogonal projector* onto the nulspace of A. That is, show that P is an orthogonal projector and $\mathrm{Ran}\,(P) = \mathrm{Nul}\,(A)$.

To see that P is an orthogonal projector, simply note that $P(I-P) = 0$, because $(A^T(AA^T)^{-1}A)^2 = A^T(AA^T)^{-1}A$. Then the orthogonality definition in 4d is satisfied.
It is obvious that $\mathrm{Nul}\,(A) \subset \mathrm{Ran}\,(P)$, since on $\mathrm{Nul}\,(A)$ the expression for P is simply the identity. To see the other direction, suppose that $w \in \mathrm{Ran}\,(P)$. Then $w = Px$ for some $x \in \mathbf{R}^n$. Then $w = x - A^T(AA^T)^{-1}Ax$. Applying A to w, we get $Aw = Ax - Ax = 0$, so $w \in \mathrm{Nul}\,(A)$.

(c) Set $h(z) = f(x_0 + Pz)$ where x_0 is any point satisfying $Ax_0 = b$. Let $\mathcal{S}_1$ be the set of first–order stationary points for the problems $\mathcal{P}$ and let $\mathcal{S}_2$ be the set of first–order stationary points for the problem $\min\{h(z) \,:\, z \in \mathbb{R}^n\}$. Show that $\mathcal{S}_1 = x_0 + P(\mathcal{S}_2)$. Show that the first-order necessary conditions for $\mathcal{P}$ are $\nabla f(x) \in \mathrm{Nul}(A)^\perp$ where $\mathrm{Nul}(A)^\perp$ is the subspace orthogonal to the null-space of A. Display both the gradient and Hessian of h.

Note that the set of solutions to $Ax = b$ is described by the set $\{x_0 + \mathrm{Nul}\,(A)\}$. Since $\mathrm{Ran}\,(P) = \mathrm{Nul}\,(A)$ by 4b), we know that $\{x_0 + \mathrm{Nul}\,(A)\} = \{x_0 + Pz\}$ where z is free to range over $\mathbf{R}^n$. For now, we will define $\mathcal{S}_1 = x_0 + P(\mathcal{S}_2)$, where $\mathcal{S}_2$ are the first-order stationary points for

$$\mathcal{P}_2: \quad \begin{array}{ll} \textit{minimize} & h(z) = f(x_0 + Pz) \\ \textit{with respect to} & z \ . \end{array}$$

The first order necessary conditions for $\mathcal{P}_2$ are simply that $P\nabla f(x_0 + Pz) = 0$. But this says that $\nabla f \in \text{Null}(P) = \text{Ran}(P)^\perp$, by the fundamental theorem of the alternative (and symmetry of P). But $\text{Ran}(P) = \text{Nul(A)}$ by 4b, so we are done.

$$\begin{aligned}\nabla h &= P\nabla f \\ \nabla^2 h &= P\nabla^2 f P\end{aligned}$$

.

(d) Show that if $\bar{x} \in I\!R^n$ is a locally optimal solution to $\mathcal{P}$, then $P\nabla f(\bar{x}) = 0$.

This follows immediately because locally optimal solutions satisfy the necessary stationarity condition given in 4c).

(e) Show that if $P\nabla f(\bar{x}) = 0$ and f is convex, the $\bar{x}$ is a globally optimal solution to $\mathcal{P}$.

Suppose that we have some point $\hat{x}$ where $f(\hat{x}) < f(\bar{x})$. Feasibility of $\hat{x}$ implies $A\hat{x} = b$, so $\lambda(\hat{x} - \bar{x}) \in \text{Nul}(A)$ for any $\lambda \in [0,1]$. Note that what we really want is for the line segment $\bar{x} + \lambda(\bar{x} - \hat{x})$ to remain feasible, and so convexity of the constrained region is enough; we do not need it to be a subspace.

By convexity of f, we have

$$f(\bar{x} + \lambda(\bar{x} - \hat{x})) \leq f(\bar{x}) + \lambda(f(\hat{x}) - f(\bar{x})).$$

But $f(\hat{x}) - f(\bar{x}) < 0$, and so we get

$$f(\bar{x} + \lambda(\bar{x} - \hat{x})) < f(\bar{x})$$

for any λ. This violates local optimality of $\bar{x}$. Note that we only used convexity of $Ax = b$, so this proof really shows that any local minimum of a convex function over a convex set is a global minimum.

6. Let $H \in I\!R^{n\times n}_s$, $u \in I\!R^n$, and $\alpha \in I\!R$ where $I\!R^{n\times n}_s$ is the linear space of all real symmetric $n \times n$ matrices. Recall that H is said to be *positive definite* if $x^T Hx > 0$ for all $x \in I\!R^n$ with $x \neq 0$. Moreover, H is said to be *positive semi–definite* if $x^T Hx \geq 0$ for all $x \in I\!R^n$. We consider the block matrix

$$\hat{H} := \begin{bmatrix} H & u \\ u^T & \alpha \end{bmatrix}.$$

(a) Show that $\hat{H}$ is positive semi–definite if and only if H is positive semi–definite and there exists a vector $z \in I\!R^n$ such that $u = Hz$ and $\alpha \geq z^T Hz$.

Suppose H is positive semidefinite, and there exists z such that $u = Hz$, and $\alpha \geq z^T Hz$. Then for any $\hat{x} = \begin{bmatrix} x \\ x_n \end{bmatrix}$ where $x_n \in \mathbf{R}$ and $x \in \mathbf{R}^{n-1}$, we have

$$\begin{aligned}\hat{x}^T \hat{H}\hat{x} &= x^T Hx + 2x^T Hx_n z + x_n^2 \alpha \\ &= (x + x_n z)^T H(x + x_n z) + x_n^2(\alpha - z^T Hz) \geq 0.\end{aligned}$$

We will prove the converse by contrapositive; i.e. we will show that if any of the conditions on H, u, or α fail, then $\hat{H}$ is not positive semidefinite.

First note that if H is not positive semidefinite then there exists some x such that $x^T Hx < 0$, and so if we take $\hat{x} = \begin{bmatrix} x \\ 0 \end{bmatrix}$ we must have $\hat{x}^T \hat{H}\hat{x} < 0$, and so $\hat{H}$ is not positive semidefinite.

Now, if we can find a vector z such that $u = Hz$ but $\alpha < z^T Hz$, take $x = -z$, and $x_n = 1$. Clearly the positive semidefinite condition fails because $\hat{x}^T \hat{H}\hat{x} = \alpha - z^T Hz < 0$.

Finally, suppose that we cannot find a vector z such that $u = Hz$. Then $u = u_1 + u_2$ where $u_1 = Hz_1$ and $u_2 \in \mathrm{Ran}(H)^\perp = \mathrm{Nul}(H)$, $u_2 \neq 0$. Then

$$\begin{aligned}\hat{x}^T \hat{H} \hat{x} &= x^T H x + 2x_n x^T (Hz_1 + u_2) + x_n^2 \alpha \\ &= (x + x_n z_1)^T H (x + x_n z_1) - 2x_n x^T u_2 + x_n^2(\alpha - z_1^T H z_1).\end{aligned}$$

Take $x_n = 1$ and $x_t = tu_2 - z_1$. Then

$$\begin{aligned}\hat{x}_t^T \hat{H} \hat{x}_t &= t^2 u_2^T H u_2 - 2tx_n(u_2 - x_n z_1)^T u_2 + x_n^2(\alpha - z_1^T H z_1) \\ &= -2t^2\|u_2\|^2 + 2tu_2^T z_1 + \alpha - z_1^T H z_1.\end{aligned}$$

Since we assumed $u_2 \neq 0$, necessarily $\|u_2\| > 0$, and so the above expression is a concave quadratic in t. It immediately follows that once t is large enough we will have $\hat{x}_t^T \hat{H} \hat{x}_t < 0$, so $\hat{H}$ is not positive semidefinite.

(b) Show that $\hat{H}$ is positive definite if and only if H is positive definite and $\alpha > u^T H^{-1} u$.

If H is positive definite and $\alpha > u^T H^u$, for any nonzero $\hat{x} = \begin{bmatrix} x \\ x_n \end{bmatrix}$ where $x_n \in \mathbf{R}$ and $x \in \mathbf{R}^{n-1}$, we have

$$\begin{aligned}\hat{x}^T \hat{H} \hat{x} &= x^T H x + 2x_n x^T H H^{-1} u + x_n^2 \alpha \\ &= (x + x_n H^{-1} u)^T H (x + x_n H^{-1}) + x_n^2(\alpha - u^T H^{-1} u) > 0,\end{aligned}$$

and so $\hat{H}$ is positive definite. To see why this is true, note that $\hat{x} \neq 0$ means either $x_n = 0$ and $x \neq 0$, or $x_n \neq 0$. In the former case, we have $(x + x_n H^{-1} u)^T H (x + x_n H^{-1}) = (x)^T H (x) > 0$ while in the latter case we have $x_n^2(\alpha - u^T H^{-1} u) > 0$.

We prove the converse by contrapositive. If H is not positive definite, we can find $x \neq 0$ such that $x^T H x \leq 0$, and then taking $\hat{x} = \begin{bmatrix} x \\ 0 \end{bmatrix}$ we get $\hat{x}^T \hat{H} \hat{x} \leq 0$, so $\hat{H}$ is not positive definite.

If $\alpha \leq u^T H^{-1} u$, take $x_n = 1$ and $x = -H^{-1} u$. Then

$$\hat{x}^T \hat{H} \hat{x} = \alpha - u^T H^{-1} u \leq 0,$$

and so $\hat{H}$ cannot be positive definite.

(c) Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $\delta \in \mathbb{R}$. Use either Part (a) or Part (b) to show that $x \in \mathbb{R}^n$ is a solution to the quadratic inequality

$$(Ax + b)^T (Ax + b) \leq c^T x + \delta$$

if and only if the block matrix

$$\begin{bmatrix} I & (Ax + b) \\ (Ax + b)^T & (c^T x + \delta) \end{bmatrix}$$

is positive semi–definite.

Take $H = I$, $u = Ax + b$, and $\alpha = c^T x + \delta$ in 6a). Take $z = u$. The conclusion of 6a) tells us that

$$(Ax + b)^T (Ax + b) \leq c^T x + \delta \iff \begin{bmatrix} I & Ax + b \\ x^T A^T + b^T & c^T x + \delta \end{bmatrix} \geq 0\,,$$

where we use notation ≥ 0 for matrices to mean 'positive semidefinite'.

(d) Suppose H is positive definite. Show that

$$\begin{bmatrix} H & u \\ 0 & (\alpha - u^T H^{-1} u) \end{bmatrix} = \begin{bmatrix} I & 0 \\ (-H^{-1} u)^T & 1 \end{bmatrix} \begin{bmatrix} H & u \\ u^T & \alpha \end{bmatrix}.$$

Perform the matrix multiplication on RHS - you get the LHS.

(e) Recall that the *kth principal minor* of a matrix $B \in I\!R^{n\times n}$ is the determinant of the upper left–hand corner k×k–submatrix of B for $1 \le k \le n$. Use an induction argument and Parts (b) and (d) above to show that H is positive definite if and only if every principal minor of H is positive.

Note: Your argument **must** use either Part (a) or Part (b) above.

Hint: det(AB)=det(A)det(B).

We will perform induction on the dimension. For $k = 1$, the statement is trivially true. Suppose we know that for $H \in \mathbf{R}^{(n-1)\times(n-1)}$ the statement holds, and consider $\hat{H} \in \mathbf{R}^{n\times n}$. Consider the decomposition in 6b. By 6b and the inductive hypothesis, $\hat{H}$ is positive definite if and only if each principal minor of H is positive and $\alpha > u^T H^{-1} u$.

By 6d) and the hint, we have $(\alpha - u^T H^{-1} u) * \det(H) = \det(\hat{H})$, so every principal minor of $\hat{H}$ is positive if and only if every principal minor of H is positive and $\alpha > u^T H^{-1} u$. But now we are done.

7. Let $A \in I\!R^{m\times n}$, $c \in I\!R^n$, $a \in I\!R^m$, $\delta > 0$, and $H \in I\!R^{n\times n}$ with H symmetric positive definite. Consider the problem

$$\begin{array}{lll} \mathcal{P} & \min_{x\in I\!R^n} c^T x & \\ & \text{subject to} & Ax = a \\ & & \|x_0 - x\|_H \le \delta \end{array}$$

where $x_0 \in I\!R^n$ satisfies $Ax_0 = a$,

$$\|z\|_H = (z^T H z)^{1/2} = [\langle z, z\rangle_H]^{1/2} ,$$

and the inner product $\langle \cdot, \cdot\rangle_H$ is defined in part (b) of problem 4 above.

(a) Suppose $H = LL^T$ for some non–singular matrix $L \in I\!R^{n\times n}$, e.g. $L = H^{1/2}$. If Q is the orthogonal projector onto the null–space of AL^{-T} in the usual (or Euclidean) inner product, show that the operator P given by

$$P = L^{-T} Q L^T$$

is the orthogonal projector onto the null–space of A with respect to the inner product $\langle \cdot, \cdot\rangle_H$.

Note that Q an orthogonal projector implies that $Q = Q^2$. Then

$$P^2 = L^{-T} Q L^T L^{-T} Q L^T = L^{-T} Q L^T = P,$$

so it is a projector. Note also that

$$HP = LL^T L^{-T} Q L^T = LQL^T = LQL^{-1} LL^T = HP^T,$$

so the criterion from 4d) is satisfied, and we know P is orthogonal with respect to $\langle \cdot, \cdot\rangle_H$.

Now we turn to the range of P. Note that $Pw = PL^{-T} L^T w = L^{-T} Q w$. Then if $w \in \mathrm{Nul}(AL^{-T})$, then $Pw = L^{-T} w$, so $APw = 0$. Then $\mathrm{Ran}(P) \subset \mathrm{Nul}(A)$. The converse is clear, since we can show that the rank of P (dimension of the the range of P) is the same as the dimension of $\mathrm{Null}(A)$, which contains it:

$$\mathrm{Rank}(P) = \mathrm{Rank}(Q) = \dim(\mathrm{Null}(AL^{-T})) = \dim(\mathrm{Null}(A)).$$

.

(b) Show that

$$\bar{x} = x_0 - \delta \|PH^{-1} c\|_H^{-1} PH^{-1} c$$

solves $\mathcal{P}$ where P is as given in part (b) above.

Hint: It may be helpful to first reduce the problem to one of the form

$$\begin{array}{ll} \min \hat{c}^T w & \\ \text{subject to} & \hat{A} w = 0 \\ & \|w\|_2^2 \le \delta^2 \ . \end{array}$$

It is also helpful to apply results relating least–squares to orthogonal projection.

Take any x_0 such that $Ax_0 = a$, and consider the minimization problem

$$\begin{array}{lll} \mathcal{P}_1 & \min_{v\in \mathbb{R}^n} c^T v & \\ & \textit{subject to} & Av = 0 \\ & & \|v\|_H^2 \le \delta^2 \end{array}$$

where $x = x_0 + v$.
Recall that any positive definite matrix H can be written $H = LL^T$, and that $\|v\|_H^2 = \|L^T v\|^2$. Also, from 7a) the matrix P is an orthogonal projector onto $\mathrm{Nul}(A)$. *Then the above problem can be reformulated again:*

$$\begin{array}{ll} \mathcal{P}_2 & \min_{w\in \mathbb{R}^n} c^T P v \\ & \textit{subject to } \|L^T P v\|^2 \le \delta^2, \end{array}$$

where Q is any orthogonal projector onto $\mathrm{Nul}(AL^{-T})$. *Since it is a projector, $Q^2 = Q$, and by criterion 4a) we also have $Q^T = Q$, so $Q = Q^T Q$.*

Plugging in for P and using the facts that $QLv = Q^T QLv$ and $L^T P v = QL^T v$, we have

$$\begin{array}{ll} \mathcal{P}_3 & \min_{v\in \mathbb{R}^n} c^T L^{-T} Q^T Q L^T v \\ & \textit{subject to } \|QL^T v\|^2 \le \delta^2 \end{array}$$

The reformulation above has the form

$$\begin{array}{lll} \mathcal{P}_4 & \min_{z\in \mathbb{R}^n} & \hat{c}^T z \\ & \textit{subject to} & AL^{-T} z = 0 \\ & & \|z\|^2 \le \delta^2 \end{array}$$

where $\hat{c} = QL^{-1}c$ and $z = QL^T v$. Note that since $\hat{c} \in$ $\mathrm{Null}(AL^{-T})$, *we can immediately write down the answer:*

$$\begin{array}{rcl} \bar{z} & = & -\delta \hat{c}\|\hat{c}\|^{-1} \\ & = & -\delta Q L^{-1} c\|QL^{-1}c\|^{-1} \\ & = & -\delta L^T L^{-T} Q L^T L^{-T} L^{-1} c\|L^T L^{-T} Q L^T L^{-T} L^{-1} c\|^{-1} \\ & = & -\delta L^T P H^{-1} c\|L^T P H^{-1} c\|^{-1} \\ & = & -\delta L^T P H^{-1} c\|P H^{-1} c\|_H^{-1} \\ & = & QL^T \bar{v}. \end{array}$$

Finally, $L^{-T}\bar{z} = P\bar{v} = -\delta P H^{-1} c\|PH^{-1}c\|_H^{-1}$, which must be the solution to $\mathcal{P}_2$, and so and $x = x_0 + P\bar{v}$ must be a solution to the original problem. Then $x = x_0 - \delta PH^{-1}c\|PH^{-1}c\|_H^{-1}$, as desired.