Exercise 10.2

(If:) Suppose that $Z^T H Z$ is positive definite. Then in particular, it is positive semidefinite, and so by Lemma 10.1, f is convex.

Also if $x, y \in \mathcal{K}$ and $x \neq y$ and $t \in (0,1)$, then as in the proof of Lemma 10.1, we have

$$(1-t) f(x) + t f(y) - f((1-t) x + ty) = \frac{1}{2} t (1-t) (y-x)^T H (y-x). \quad (*)$$

But $x - y \in \ker A$ and so $x - y = Zv$ for some $v \in \mathbb{R}^k$.

Hence $(*)$ becomes

$$(1-t) f(t) + t f(y) - f((1-t) x + ty) = \frac{1}{2} t (1-t) v^T Z^T H Z v. \quad (**)$$

Also $x \neq y$ and so $v \neq 0$ (otherwise $x - y = Z0 = 0$ and so $x = y$!). Hence by the positive definiteness of $Z^T H Z$, $v^T Z^T H Z v > 0$.

Consequently $(1-t) f(t) + t f(y) > f((1-t) x + ty)$ (from $(**)$)

So f is strictly convex.

(Only if:) Suppose f is strictly convex. By Lemma 10.1, $Z^T H Z$ is positive semi-definite. Let $\bar{x} \in \mathbb{R}^n$ be such that $A\bar{x} = b$. Take $x = \bar{x}$ and $y = \bar{x} + Zv$, where $v \in \mathbb{R}^k$ and $v \neq 0$. Let $t = \frac{1}{2}$. Then we have

$$0 < (1-t) f(x) + t f(y) - f((1-t) x + ty)$$
$$= \frac{1}{2} (t - t^2) (y-x)^T H (y-x) = \frac{1}{8} v^T Z^T H Z v,$$

and so $v^T Z^T H Z v > 0$. So $Z^T H Z$ is positive definite

Exercise 10.7

An equivalent quadratic optimization problem is:

$$(Q): \begin{cases} \text{minimize} & (x-y)^T(x-y), \\ \text{subject to} & a^T x = b. \end{cases}$$

We observe that

$$(x-y)^T(x-y) = x^T x - 2y^T x + y^T y$$

and so (Q) is the quadratic optimization problem

$$(Q): \begin{cases} \text{minimize} & \frac{1}{2} x^T H x + c^T x + c_0 \\ \text{subject to} & Ax = b, \end{cases}$$

where

$$H = 2I,$$
$$c = -2y,$$
$$c_0 = y^T y,$$
$$A = a^T,$$
$$b = b.$$

$\hat{x}$ is an optimal solution iff:

(1) $A\hat{x} = b$, and

(2) $\exists\ u \in \mathbb{R}^m$ such that $H\hat{x} + c = A^T u$.

In our special case above, we obtain

(1) $a^T \hat{x} = b$

(2) $2I\hat{x} - 2y = ua$, i.e., $2\hat{x} = ua + 2y$.

By premultiplying (2) by a^T, we have using (1):

$$2b = 2a^T\hat{x} = u\, a^T a + 2a^T y = u\|a\|^2 + 2a^T y,$$

and so $u = \dfrac{2(b - a^T y)}{\|a\|^2}$.

From (2) we now obtain that

$$\hat{x} = \frac{1}{2} ua + y = \frac{1}{2} \cdot \frac{2(b - a^T y)}{\|a\|^2} \cdot a + y$$

$$= \frac{b - a^T y}{\|a\|^2} \cdot a + y.$$

Hence $\hat{x} - y = \dfrac{b - a^T y}{\|a\|^2} \cdot a$, and so $\|\hat{x} - y\| = \dfrac{|b - a^T y|}{\|a\|}$.

Exercise 10.8

We want to solve $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 10 \\ 14 \end{bmatrix}$.

If we set $\gamma = 0$, we obtain $\begin{Bmatrix} \alpha + 2\beta = 10 \\ 3\alpha + 2\beta = 14 \end{Bmatrix}$, which has the solution $\alpha = 2$, $\beta = 4$.

So one solution to $Ax = b$ is given by $\bar{x} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$.

ker A: $x \in \ker A$ iff $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$ i.e., $\underbrace{\begin{cases} x_1 + 2x_2 = -3 \\ 3x_1 + 2x_2 = -x \end{cases}}_{(*)}$

But (*) has the solution $x_1 = x_3$ and $x_2 = -2x_3$.

Consequently $\ker A = \operatorname{span} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \operatorname{ran} Z$, where $Z = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

We want to find an optimal solution to the quadratic optimization problem

$$(Q): \begin{cases} \text{minimize} & f(x) \quad (= \tfrac{1}{2} x^T H x) \\ \text{subject to} & Ax = b \end{cases}$$

where $H = \begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix}$.

We know that $Ax = b$ iff $x = \bar{x} + Zv$ for $\overbrace{\text{}}^{\text{some}}$ $v \in \mathbb{R}$.

We have

$$Z^T H Z = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 & -12 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 36,$$

$$Z^T H \bar{x} = \begin{bmatrix} 6 & -12 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} = 12 - 48 = -36.$$

Hence the optimal solution to (Q) is given by

$\hat{x} = \bar{x} + Z\hat{v}$, where $\hat{v}$ satisfies $(Z^T H Z)\hat{v} = -Z^T(H\bar{x}+$

So

$36\hat{v} = +36$ and so $\hat{v} = +1$.

Hence $\hat{x} = \bar{x} + Z\hat{v} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

$= \begin{bmatrix} 3 \\ 2 \\ +1 \end{bmatrix}$.

Exercise 10.9

(1) $A\bar{x} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = b$, and so

$\bar{x}$ is a feasible solution.

(2) Since rank $A = 3$, $\dim(\ker A) = 5 - 3 = 2$.

z_1 and z_2 are linearly independent, and they belong to $\ker A$:

$$A z_1 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$A z_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1-1+0 \\ 0+0+0 \\ 0-1+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Hence $\ker A = \text{span}\{z_1, z_2\}$.

So $\{z_1, z_2\}$ forms a basis for $\ker A$.

(3) $$Z^T H Z = \begin{bmatrix} 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ 2 & 0 & -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$$

$$Z^T H \bar{x} = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ 2 & 0 & -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$(Z^T H Z)\hat{v} = -Z^T(H\bar{x} + c)$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \hat{v} = -\begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

So $\hat{v} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{3} \end{bmatrix}$.

Hence

$\hat{x} = \bar{x} + Z\hat{v}$, and so $\left.\begin{aligned} \hat{x} = {} & \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T \\ & + \begin{bmatrix} -\frac{1}{3} & 0 & +\frac{1}{3} & 0 & -\frac{1}{3} \end{bmatrix}^T \end{aligned}\right\} = \begin{bmatrix} \frac{2}{3} & 1 & \frac{4}{3} & 1 & \frac{2}{3} \end{bmatrix}$

Exercise 10.10

(1) We have $\|x-q\|^2 = (x-q)(x-q) = x^Tx - 2q^Tx + q^Tq$
$$= \frac{1}{2}x^THx + c^Tx + c_0,$$
where
$$H = 2I, \quad c = -2q \quad \text{and} \quad c_0 = q^Tq.$$
The constraint is that $x \in \ker A = \{x : Ax = 0\}$.
Using the Lagrange method, we know that $\hat{x}$ is optimal iff $\exists u$ such that
$$\begin{bmatrix} H & -A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \hat{x} \\ u \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix},$$
and in our case this gives
$$2I\hat{x} - A^Tu = +2q$$
$$A\hat{x} = 0.$$
Thus $\hat{x} = q + \frac{1}{2}A^Tu$ and $A\hat{x} = 0$.

So $0 = A\hat{x} = Aq + \frac{1}{2}AA^Tu$.
We have
$$Aq = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix},$$
$$\frac{1}{2}AA^T = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Consequently $u = -\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1}\begin{bmatrix} 8 \\ 4 \end{bmatrix} = -\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}.$

Hence
$$\hat{x} = q + \frac{1}{2}A^Tu = \begin{bmatrix} 4 \\ 2 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \\ -2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

(2) $x \in \operatorname{ran} A^T$ iff $x = A^T v$ for some v.

So the problem can be rephrased as follows:

$$(*) \begin{cases} \text{minimize} & \|A^T v - q\|^2 \\ \text{subject to} & v \in \mathbb{R}^2 . \end{cases}$$

But

$$\begin{aligned} \|A^T v - q\|^2 &= (A^T v - q)^T (A^T v - q) \\ &= v^T A A^T v - 2\, q^T A^T v + q^T q \\ &= \tfrac{1}{2} v^T H v + c^T v + c_0 , \end{aligned}$$

where $H = 2AA^T$, $c = -2Aq$ and $c_0 = q^T q$.

$\hat{v}$ is optimal for $(*)$ iff

$$H\hat{v} = -c$$

i.e., $2AA^T\hat{v} = +2Aq$

i.e., $AA^T\hat{v} = Aq$

i.e., $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \hat{v} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$.

Thus

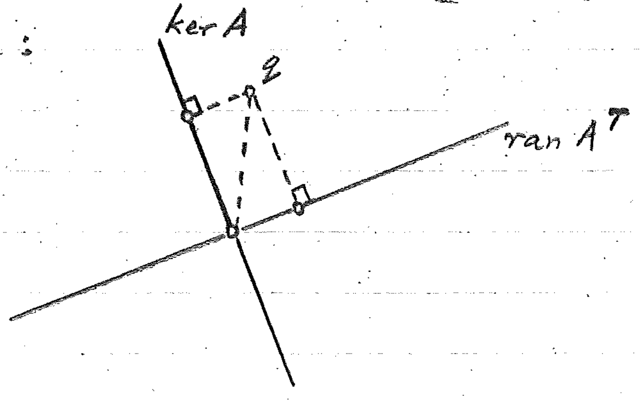
$$\hat{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

and so

$$\hat{x} = A^T\hat{v} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

is optimal for (P2).

(We observe that the optimal solutions to (P1) and (P2) are orthogonal and their sum is q. This is not surprising since $\operatorname{ran} A^T = (\ker A)^{\perp}$; see the following figure:



)