Exercise 9.8.

(1). $f(x) = x_1^2 + 2x_2^2 + 5x_3^2 + 3x_2x_3$

$$= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & \frac{3}{2} \\ 0 & \frac{3}{2} & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x^T H x,$$

where $H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & \frac{3}{2} \\ 0 & \frac{3}{2} & 5 \end{bmatrix}$.

Since

$$f(x) = x^T H x = x_1^2 + \left(\sqrt{2}\, x_2 + \frac{3\sqrt{2}}{4} x_3\right)^2 + 5x_3^2 - \frac{9 \cdot 2}{16} x_3^2$$

$$= x_1^2 + \left(\sqrt{2}\, x_2 + \frac{3\sqrt{2}}{4} x_3\right)^2 + \frac{31}{8} x_3^2 \geq 0 \quad \forall x$$

we see that f is convex since H is positive semi-definite.

In fact if $x^T H x = 0$, then $x_1 = x_3 = x_2 = 0$, and so H is positive definite. So f is strictly convex

(2). $f(x) = 2x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 + 2x_1x_3$

$= x^T H x,$

where $H = \begin{bmatrix} 2 & -1 & +1 \\ -1 & 1 & 0 \\ +1 & 0 & 1 \end{bmatrix}$.

$$x^T H x = 2x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 + 2x_1x_3$$

$$= x_1^2 - 2x_1x_2 + x_2^2 + x_1^2 + 2x_1x_3 + x_3^2$$

$$= (x_1 - x_2)^2 + (x_1 + x_3)^2 \geq 0 \quad \forall x$$

and so H is positive semi-definite. Hence f is convex.

H is not positive definite, since for example with $x = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, we have $x^T H x = (1-1)^2 + (1+-1)^2 = 0^2 + 0^2 = 0$, but $x \neq 0$. So f is not strictly convex.

Exercise 9.9

$f(x) = x_1^2 + 2x_2^2 + 2a\,x_1 x_2 = x^T H x,$

where

$$H = \begin{bmatrix} 1 & a \\ a & 2 \end{bmatrix}.$$

H is positive semi-definite iff all eigenvalues of H are nonnegative.

H is positive definite iff all eigenvalues of H are positive.

Eigenvalues of H:

$$\det \begin{bmatrix} \lambda-1 & -a \\ -a & \lambda-2 \end{bmatrix} = (\lambda-1)(\lambda-2) - a^2 = \lambda^2 - 3\lambda + 2 - a^2 = 0$$

$$\text{iff } \lambda = \frac{3 \pm \sqrt{9-4(2-a^2)}}{2} = \frac{3 \pm \sqrt{1+a^2}}{2}.$$

$3 + \sqrt{1+a^2} > 0 \quad \forall a \in \mathbb{R}.$

$3 - \sqrt{1+a^2} \geq 0 \quad \text{iff} \quad 9 \geq 1+a^2 \quad \text{iff} \quad -2\sqrt{2} \leq a \leq 2\sqrt{2}.$

Thus

(1) f is convex iff $-2\sqrt{2} \leq a \leq 2\sqrt{2}$.

(2) f is strictly convex iff $-2\sqrt{2} < a < 2\sqrt{2}$.

Exercise 9.12

Suppose d is a descent direction at $x \in \mathbb{R}^n$ and $\forall t \in (0, \epsilon)$, $f(x+td) \leq f(x)$. Suppose $(Hx+c)^T d > 0$. Then we have

$$f(x+td) = f(x) + t(Hx+c)^T d + \frac{1}{2} t^2 d^T H d$$
$$= f(x) + \frac{1}{2} t \left[2(Hx+c)^T d + t d^T H d \right]$$
$$> f(x)$$

for all $t > 0$ if $d^T H d \geqslant 0$, and for all $0 < t < \frac{\delta}{(-d^T H d)}$ if $d^T H d < 0$. This contradicts the fact that d is a descent direction at $x \in \mathbb{R}^n$. Hence $(Hx+c)^T d \leq 0$

Exercise 9.15

We have

$$(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2 = 2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_2x_3-2x_3x$$

$$= \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \frac{1}{2} x^T H x,$$

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

H is positive semi-definite, since

$$\frac{1}{2} x^T H x = (x_1-x_2)^2 + (x_2-x_3)^2 + (x_3-x_1)^2 \geqslant 0 \quad \forall\, x \in \mathbb{R}^3.$$

Kernel of H: $\ker H = \{x \in \mathbb{R}^3 : Hx = 0\}$.

$$Hx = 0 \iff \begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix} x = 0$$

$$\underset{\frac{1}{2}\cdot\text{row }1}{\iff} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix} x = 0$$

$$\underset{\substack{\underline{\text{row }2} + 2\cdot\text{row }1 \\ \underline{\text{row }3} + 2\cdot\text{row }1}}{\iff} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 3 & -3 \\ 0 & -3 & 3 \end{bmatrix} x = 0$$

$$\underset{\frac{1}{3}\cdot\text{row }2}{\iff} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & -3 & 3 \end{bmatrix} x = 0$$

$$\Longleftrightarrow \quad \begin{array}{l} \underline{\text{row } 3} + 3\,\text{row } 2 \\ \underline{\text{row } 1} + \tfrac{1}{2}\,\text{row } 2 \end{array} \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} x = 0$$

$$\Leftrightarrow \quad x_1 = x_2 = x_3 \quad \Leftrightarrow \quad x \in \operatorname{span} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

So $\ker H = \operatorname{span} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

x is a minimizer of f iff $Hx = -c$.

So f has at least one minimizer iff $-c \in \operatorname{ran} H$.

But $\operatorname{ran} H = (\ker H^T)^\perp = (\ker H)^\perp = \left(\operatorname{span} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)^\perp$.

So $-c \in \operatorname{ran} H$ iff $c \in \left(\operatorname{span} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)^\perp$ i.e.,

$$[1 \;\; 1 \;\; 1]\, c = 0, \text{ i.e., } v^T c = 0,$$

where $v := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

For example if $c = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, then $v^T c = 0$ and so in this case f has a minimizer.

But if $c = v = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, then $v^T c = v^T v = 3 \neq 0$ and so in this case $-c \notin \operatorname{ran} H$. Consequently, f is not bounded below (see Theorem 9.14).

Exercise 9.16.

(1) Let $x(\alpha) = a + \alpha \cdot u \in L_1$ and $y(\beta) = b + \beta \cdot v \in L_2$.

The square of the distance between $x(\alpha)$ and $y(\beta)$ is

$$f(\alpha,\beta) := \| (a+\alpha \cdot u) - (b+\beta v) \|_2^2$$

and the problem is that of minimizing f.

We have

$$\begin{aligned} f(\alpha,\beta) &= ((a+\alpha u) - (b+\beta v))^T ((a+\alpha u) - (b+\beta v))^T \\ &= (a-b+\alpha u - \beta v)^T (a-b+\alpha u - \beta v) \\ &= (a-b)^T(a-b) + \underbrace{u^T u}_{=1} \alpha^2 + \underbrace{v^T v}_{=1} \beta^2 - 2u^T v\, \alpha\beta + \alpha\, u^T(a-b) \\ &\qquad + \beta\, v^T(b-a) \\ &= \alpha^2 + \beta^2 - 2u^T v\, \alpha\beta + 2u^T(a-b)\alpha + 2v^T(b-a)\beta + (a-b)^T(a-b) \\ &= \tfrac{1}{2} \xi^T H \xi + c^T \xi + c_0, \end{aligned}$$

where $\xi := \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, $c := \begin{bmatrix} 2u^T(a-b) \\ 2v^T(b-a) \end{bmatrix}$, $c_0 := (a-b)^T(a-b)$

and $H = 2 \begin{bmatrix} 1 & -u^T v \\ -u^T v & 1 \end{bmatrix}$.

(2) H is symmetric, and its eigenvalues are determined by

$$\det \begin{bmatrix} \lambda - 1 & -u^T v \\ -u^T v & \lambda - 1 \end{bmatrix} = 0 \quad \text{i.e.,} \quad (\lambda-1)^2 - (u^T v)^2 = 0$$

$$\text{i.e.,} \quad \lambda = 1 \pm u^T v \geq 0$$

and so the eigenvalues of H are $2(1+u^T v)$ and $2(1-u^T v)$ which are positive. So H is strictly convex.

(3) The optimal $\hat{\xi}$ is determined by the equation

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

and so

$$2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \hat{\xi} = \begin{bmatrix} 2u^T(b-a) \\ 2v^T(a-b) \end{bmatrix}.$$

Hence $\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \hat{\xi} = \begin{bmatrix} u^T(b-a) \\ v^T(a-b) \end{bmatrix}$. Hence $\hat{x} = a + \hat{\alpha} u = a + (u^T(b-a))u$

and

$\hat{y} = b + \hat{\beta} v = b + (v^T(a-b))$