Non-linear Optimisation Answers

1. $\nabla f_1(\mathbf{x}) = \begin{pmatrix} 2x_1 - 3x_2 \\ -3x_1 + 6x_2 \end{pmatrix}$, $\mathbf{H}f_1(\mathbf{x}) = \begin{pmatrix} 2 & -3 \\ -3 & 6 \end{pmatrix}$.
Stationary point when $2x_1 - 3x_2 = -3x_1 + 4x_2 = 0$, i.e., $x_1 = x_2 = 0$.
$\det \mathbf{H}f_1(\mathbf{0}) = 3 > 0$, leading coefficient $2 > 0$ so this is a local minimum.

$\nabla f_2(\mathbf{x}) = \begin{pmatrix} 4x_1 - 5x_2 \\ -5x_1 + 4x_2 \end{pmatrix}$, $\mathbf{H}f_2(\mathbf{x}) = \begin{pmatrix} 4 & -5 \\ -5 & 4 \end{pmatrix}$.
Stationary point when $4x_1 - 5x_2 = -5x_1 + 4x_2 = 0$, i.e., $x_1 = x_2 = 0$.
$\det \mathbf{H}f_2(\mathbf{0}) = -9 < 0$ so this is a saddle point.

$f_3(\mathbf{x}) = x_1x_2 - x_1^2x_2 - x_1x_2^2$,

$$\nabla f_3(\mathbf{x}) = \begin{pmatrix} x_2 - 2x_1x_2 - x_2^2 \\ x_1 - x_1^2 - 2x_1x_2 \end{pmatrix} = \begin{pmatrix} x_2(1 - 2x_1 - x_2) \\ x_1(1 - x_1 - 2x_2) \end{pmatrix},$$

$$\mathbf{H}f_3(\mathbf{x}) = \begin{pmatrix} -2x_2 & 1 - 2x_1 - 2x_2 \\ 1 - 2x_1 - 2x_2 & -2x_1 \end{pmatrix}.$$

Stationary points when
either $x_2 = 0$, $x_1 = 0$
or $x_2 = 0$, $1 - x_1 - 2x_2 = 0$, i.e., $x_1 = 1$, $x_2 = 0$
or $1 - 2x_1 - x_2 = 0$, $x_1 = 0$, i.e., $x_1 = 0$, $x_2 = 1$
or $1 - 2x_1 - x_2 = 0$, $1 - x_1 - 2x_2 = 0$, i.e., $x_1 = x_2 = \frac{1}{3}$.

$\mathbf{H}f_3(\mathbf{0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has determinant $-1 < 0$ so this is a saddle point.

$\mathbf{H}f_3\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & -2 \end{pmatrix}$ has determinant $-1 < 0$ so this is a saddle point.

$\mathbf{H}f_3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix}$ has determinant $-1 < 0$ so this is a saddle point.

$\mathbf{H}f_3\begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{2}{3} \end{pmatrix}$ has determinant $\frac{1}{3} > 0$ and leading coefficient $-\frac{2}{3} < 0$ so this is a local maximum.

$\nabla f_4(\mathbf{x}) = \begin{pmatrix} \cos(x_1 + x_2) \\ \cos(x_1 + x_2) \end{pmatrix}$.

$\mathbf{H}f_4(\mathbf{x}) = \begin{pmatrix} -\sin(x_1 + x_2) & -\sin(x_1 + x_2) \\ -\sin(x_1 + x_2) & -\sin(x_1 + x_2) \end{pmatrix}$.

Stationary points when $\cos(x_1 + x_2) = 0$, i.e., $x_1 + x_2$ is an odd multiple of $\frac{\pi}{2}$.

At such stationary points, either $f_4(\mathbf{x}) = 1$ (when $x_1 + x_2 = 2k\pi + \frac{\pi}{2}$ for some integer k) and $\mathbf{x}$ is a local maximum, or $f_4(\mathbf{x}) = -1$ (when $x_1 + x_2 = 2k\pi - \frac{\pi}{2}$ for some integer k) and $\mathbf{x}$ is a local minimum.

(The determinant of $\mathbf{H}f_4$ is always 0, so tells us nothing.)

$$\nabla f_5(\mathbf{x}) = \begin{pmatrix} 4x_1 + 2x_2 + 2x_3 \\ 2x_1 + 6x_2 + 2x_3 \\ 2x_1 + 2x_2 + 4x_3 \end{pmatrix} = 2\begin{pmatrix} 2x_1 + x_2 + x_3 \\ x_1 + 3x_2 + x_3 \\ x_1 + x_2 + 2x_3 \end{pmatrix}.$$

$\mathbf{H}f_5(\mathbf{x}) = 2\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Stationary point when $2x_1 + x_2 + x_3 = x_1 + 3x_2 + x_3 = x_1 + x_2 + 2x_3 = 0$, i.e., at $\mathbf{0}$.

$\mathbf{H}f_5(\mathbf{0})$ has determinant $2^3 \cdot 7 > 0$; its principal minor $2\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ has determinant $2^2 \cdot 5 > 0$; and its leading coefficient is $2 \cdot 2 > 0$. So $\mathbf{0}$ is a local minimum.

(Alternative method: observe that $f_5(\mathbf{x}) = (x_1 + x_2)^2 + (x_2 + x_3)^2 + (x_1 + x_3)^2 + x_2^2$, so is zero at $\mathbf{0}$ and strictly positive everywhere else.)

$$\nabla f_6(\mathbf{x}) = \begin{pmatrix} -2x_1x_3e^{-x_1^2-x_2^2} \\ -2x_2x_3e^{-x_1^2-x_2^2} \\ e^{-x_1^2-x_2^2} \end{pmatrix} = e^{-x_1^2-x_2^2}\begin{pmatrix} -2x_1x_3 \\ -2x_2x_3 \\ 1 \end{pmatrix}.$$

$$\begin{aligned}\mathbf{H}f_6(\mathbf{x}) &= \begin{pmatrix} (4x_1^2x_3-2x_3)e^{-x_1^2-x_2^2} & 4x_1x_2x_3e^{-x_1^2-x_2^2} & -2x_1e^{-x_1^2-x_2^2} \\ 4x_1x_2x_3e^{-x_1^2-x_2^2} & (4x_2^2x_3-2x_3)e^{-x_1^2-x_2^2} & -2x_2e^{-x_1^2-x_2^2} \\ -2x_1e^{-x_1^2-x_2^2} & -2x_2e^{-x_1^2-x_2^2} & 0 \end{pmatrix} \\ &= 2e^{-x_1^2-x_2^2}\begin{pmatrix} 2x_1^2x_3-x_3 & 2x_1x_2x_3 & -x_1 \\ 2x_1x_2x_3 & 2x_2^2x_3-x_3 & -x_2 \\ -x_1 & -x_2 & 0 \end{pmatrix}.\end{aligned}$$

No stationary points, because $\frac{\partial f_6}{\partial x_3}$ is never 0.

2. (a) f_1 is not coercive. There are many ways of seeing this. (i) We can express it as $(x_1-2x_2)(x_1-x_2)$, so that it is zero if either $x_1 = 2x_2$ or $x_1 = x_2$, and either can be true for large $\|\mathbf{x}\|$. (ii) Its only stationary point is at $\mathbf{0}$, and the Hessian there is $\begin{pmatrix} 1 & -3 \\ -3 & 4 \end{pmatrix}$, which has determinant less than 0. This is therefore a saddle point, and f_1 has no local minimum, let alone a global minimum.

(b) f_2 is coercive, because it is $(x_1-x_2)^2+2x_2^2$, which is large if either $|x_2|$ is large or if x_2 is small and x_1 is large.

(c) f_3 is not coercive; if we set $x_1 = x_3 = 0$ and $x_2 < 0$ we can get $f_3(\mathbf{x}) < 0$ while $\|\mathbf{x}\| = |x_2| = -x_2$ is as large as we like.

(d) f_4 is not coercive, because if $x_1 = x_2 = x_3 = \alpha$ then $f_4(\mathbf{x}) = -\alpha^4$.

(e) f_5 is coercive, because all the terms $\cos x_2x_3$, etc., are between -1 and 1, so that if $\alpha \geq 1$ and $\|\mathbf{x}\| \geq (\alpha+1)\sqrt{3}$ (so that at least one of $|x_1|$, $|x_2|$, $|x_3|$ is greater than or equal to $\alpha+1$) then $f(\mathbf{x}) \geq \alpha^2$.

3. $f''(x) = 2+6cx+12x^2 = 12(x+\frac{c}{4})^2+(2-\frac{3c^2}{4})$, so f is convex iff $2 \geq \frac{3c^2}{4}$, that is, $c^2 \leq \frac{8}{3}$, that is, $|c| \leq \sqrt{\frac{8}{3}}$.

4. (i) $\mathbf{H}f = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, which is positive definite (its principal minors have determinants 2, 3 and 4), so f is convex.

(ii) $\mathbf{H}f = \begin{pmatrix} 12x_1^2 & 3x_2^2 \\ 3x_2^2 & 6x_1x_2+12x_2^2 \end{pmatrix}$ has determinant $144x_1^2x_2^2+72x_1^3x_2-9x_2^4$ which is sometimes negative, so $\mathbf{H}f$ cannot always be positive semi-definite and f is not convex.

(iii) $\mathbf{H}f = e^{x_1-x_2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, which is positive semi-definite (its eigenvalues are 0 and $2e^{x_1-x_2}$), so f is convex. [*Note to mathematicians*: what is happening here is that $f = g \circ h$, where $h(\mathbf{x}) = x_1 - x_2$ and $g(x) = e^x$ are convex *and g is monotonic increasing.*]

5. (i) $\mathbf{H}f = 4e^{x_1+x_2}(e^{x_1+x_2}-1)\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$ has eigenvalues 0 and $4e^{x_1+x_2}(e^{x_1+x_2}-1)$; the latter can be less than 0, so the matrix is not always positive semi-definite, and the function is not convex. [*Note to mathematicians*: the functions $\mathbf{x} \mapsto e^{x_1+x_2}$, $x \mapsto (x-2)^2$ are convex, but the latter is not monotonic increasing.]

(ii) $\mathbf{H}f = \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 1 & 2\end{pmatrix}$. [This is not positive definite; its principal minors have determinants 1, 1 and 0.] Its eigenvalues are given by the equation

$$\begin{aligned} 0 &= \det\begin{pmatrix}1-\lambda & 0 & 1\\ 0 & 1-\lambda & 1\\ 1 & 1 & 2-\lambda\end{pmatrix} \\ &= (1-\lambda)(1-\lambda)(2-\lambda) - (1-\lambda) - (1-\lambda) = (1-\lambda)(2-3\lambda+2\lambda^2-2) \\ &= -\lambda(\lambda-1)(\lambda-3), \end{aligned}$$

so that the eigenvalues 0, 1 and 3 are all non-negative, $\mathbf{H}f$ is positive semi-definite and f is convex.

6. (a)(i) Yes. If $\mathbf{x},\ \mathbf{y} \in C_1 \cap C_2$ and $t \in [0,1]$ then

$$\mathbf{x},\ \mathbf{y} \in C_1, \text{ which is convex, so } t\mathbf{x} + (1-t)\mathbf{y} \in C_1,$$

$$\mathbf{x},\ \mathbf{y} \in C_2, \text{ which is convex, so } t\mathbf{x} + (1-t)\mathbf{y} \in C_2,$$

and $t\mathbf{x} + (1-t)\mathbf{y} \in C_1 \cap C_2$; as $\mathbf{x}$, $\mathbf{y}$ and t are arbitrary, $C_1 \cap C_2$ is convex.

(ii) No. For example, try $C_1 = [0,1]$, $C_2 = [2,3]$; these are both convex subsets of $\mathbb{R}$ but their union is not convex.

(b)(i) Yes. If $\mathbf{x},\ \mathbf{y} \in C$ and $t \in [0,1]$, then

$$\begin{aligned}(f+g)(t\mathbf{x} + (1-t)\mathbf{y}) &= f(t\mathbf{x} + (1-t)\mathbf{y}) + g(t\mathbf{x} + (1-t)\mathbf{y})\\ &\le tf(\mathbf{x}) + (1-t)f(\mathbf{y}) + tg(\mathbf{x}) + (1-t)g(\mathbf{y})\end{aligned}$$

(because f and g are convex)

$$= t(f+g)(\mathbf{x}) + (1-t)(f+g)(\mathbf{y});$$

as $\mathbf{x}$, $\mathbf{y}$ and t are arbitrary, $f+g$ is convex.

(ii) No. For instance, if we set $C = \mathbb{R}$, $f(x) = x^2$, $a = -1$ then af is not convex.

(iii) Yes. If $\mathbf{x},\ \mathbf{y} \in C$ and $t \in [0,1]$, then
— if $f(t\mathbf{x} + (1-t)\mathbf{y}) \le g(t\mathbf{x} + (1-t)\mathbf{y})$, then

$$\max(f,g)(t\mathbf{x} + (1-t)\mathbf{y}) = g(t\mathbf{x} + (1-t)\mathbf{y}) \le tg(\mathbf{x}) + (1-t)g(\mathbf{y})$$

(because g is convex)

$$\le t\max(f,g)(\mathbf{x}) + (1-t)\max(f,g)(\mathbf{y});$$

— if $g(t\mathbf{x} + (1-t)\mathbf{y}) \le f(t\mathbf{x} + (1-t)\mathbf{y})$, then

$$\max(f,g)(t\mathbf{x} + (1-t)\mathbf{y}) = f(t\mathbf{x} + (1-t)\mathbf{y}) \le tf(\mathbf{x}) + (1-t)f(\mathbf{y})$$

(because f is convex)

$$\le t\max(f,g)(\mathbf{x}) + (1-t)\max(f,g)(\mathbf{y});$$

thus $\max(f,g)(t\mathbf{x}+(1-t)\mathbf{y}) \le t\max(f,g)(\mathbf{x}) + (1-t)\max(f,g)(\mathbf{y})$ in both cases. As $\mathbf{x}$, $\mathbf{y}$ and t are arbitrary, $\max(f,g)$ is convex.

(iv) No. For example, if $C = \mathbb{R}$, $f(x) = e^x$, $g(x) = e^{-x}$ then $\min(f,g)(x) = e^{-|x|}$ is not convex.

(v) No. For example, if $C = \mathbb{R}$, $f(x) = g(x) = x^2 - 1$ then $f \times g$ is not convex (it takes a greater value at 0 than at either ± 1.)

7. (a) Continuous; near the transition at $x = 0$, both $(x-1)^2$ and e^x are approximately 1.

(b) Discontinuous; near the transition at $x = 0$, $f_2(x) \doteq 0$ if $x > 0$ and $f_2(x) \doteq e$ if $x < 0$.

(c) Discontinuous; near the transition point $(0,1)$, $g_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \doteq 0$ if $x_1 > 0$ but $g_1\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \doteq \sin 1$ if $x_1 < 0$.

(d) Continuous; near any transition point (x,x), both $\sin x_1 \cos x_2$ and $\cos x_1 \sin x_2$ are approximately $\sin x \cos x$.

8. (a) $\dfrac{\partial f}{\partial x_1} = e^{-x_2}, \quad \dfrac{\partial f}{\partial x_2} = -x_1 e^{-x_2},$

$$\frac{\partial^2 f}{\partial x_1^2} = 0, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -e^{-x_2}, \quad \frac{\partial^2 f}{\partial x_2^2} = x_1 e^{-x_2}.$$

(b) $\dfrac{\partial g}{\partial x_1} = x_2^2 x_3^3 + 3x_2 x_3^2 x_1^2 + 2x_3 x_1 x_2^3, \quad \dfrac{\partial g}{\partial x_2} = x_3^2 x_1^3 + 3x_3 x_1^2 x_2^2 + 2x_1 x_2 x_3^3,$

$$\frac{\partial g}{\partial x_3} = x_1^2 x_2^3 + 3x_1 x_2^2 x_3^2 + 2x_2 x_3 x_1^3,$$

$$\frac{\partial^2 g}{\partial x_1^2} = 6x_2 x_3^2 x_1 + 2x_3 x_2^3, \quad \frac{\partial^2 g}{\partial x_2^2} = 6x_3 x_1^2 x_2 + 2x_1 x_3^3, \quad \frac{\partial^2 g}{\partial x_3^2} = 6x_1 x_2^2 x_3 + 2x_2 x_1^3,$$

$$\frac{\partial^2 g}{\partial x_1 \partial x_2} = 2x_2 x_3^3 + 3x_3^2 x_1^2 + 6x_3 x_1 x_2^2 = \frac{\partial^2 g}{\partial x_2 \partial x_1},$$

$$\frac{\partial^2 g}{\partial x_2 \partial x_3} = 2x_3 x_1^3 + 3x_1^2 x_2^2 + 6x_1 x_2 x_3^2 = \frac{\partial^2 g}{\partial x_3 \partial x_2},$$

$$\frac{\partial^2 g}{\partial x_3 \partial x_1} = 2x_1 x_2^3 + 3x_2^2 x_3^2 + 6x_2 x_3 x_1^2 = \frac{\partial^2 g}{\partial x_1 \partial x_3}.$$

9. (i) $1 \cdot 3 - 2 \cdot 2 = -1$.

(ii) $1 \cdot 3 \cdot 5 + 2 \cdot 4 \cdot 3 + 3 \cdot 2 \cdot 4 - 1 \cdot 4 \cdot 4 - 2 \cdot 2 \cdot 5 - 3 \cdot 3 \cdot 3 = 0$. (Or note that the top and bottom rows sum to twice the middle row.)

(iii) $0 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 - 0 \cdot 1 \cdot 1 - 1 \cdot 1 \cdot 0 - 1 \cdot 0 \cdot 1 = 2$.

10. (a) None of these. [$\det \mathbf{A} < 0$ and dimension is even, so one eigenvalue must be positive and the other negative.]

(b) Positive definite. [$\det \mathbf{A} = 2 \cdot 1 \cdot 2 - 2 \cdot 1 \cdot 1 - 1 \cdot 1 \cdot 1 = 1 > 0$; also $\det(\,2\,) = \det \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2 > 0$ so all eigenvalues are greater than 0.]

(c) Negative definite. [$\det \mathbf{A} = 3$ so either both are positive or both are negative. But the principal minor $\det(\,-2\,)$ is negative, so they cannot both be positive and must both be negative.]

(d) None of these. [$\det \mathbf{A} > 0$ (see question 9(iii)), so no eigenvalue can be 0, and also (because the dimension is odd) they cannot all be negative; on the other hand, the principal minors $\det(\,0\,) = 0$ and $\det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1$ are not both positive, so not all the eigenvalues of $\mathbf{A}$ can be greater than 0.]

(e) None of these. [$\det \mathbf{A} = 0$ (see question 9(ii)), so 0 is an eigenvalue. To determine the other eigenvalues, observe that

$$\begin{aligned}
&\det \begin{pmatrix} 1-\lambda & 2 & 3 \\ 2 & 3-\lambda & 4 \\ 3 & 4 & 5-\lambda \end{pmatrix} \\
&= (1-\lambda)((3-\lambda)(5-\lambda) - 16) - 2(2(5-\lambda) - 12) + 3(8 - 3(3-\lambda)) \\
&= (1-\lambda)(15 - 8\lambda + \lambda^2 - 16) - 2(10 - 2\lambda - 12) + 3(8 - 9 + 3\lambda) \\
&= (1-\lambda)(-1 - 8\lambda + \lambda^2) - 2(-2\lambda - 2) + 3(-1 + 3\lambda) \\
&= -1 - 7\lambda - 7\lambda^2 - \lambda^3 + 4\lambda + 4 - 3 + 9\lambda \\
&= 6\lambda + 9\lambda^2 - \lambda^3 = \lambda(6 + 9\lambda - \lambda^2)
\end{aligned}$$

is zero iff either $\lambda = 0$ or $\lambda^2 - 9\lambda - 6 = 0$ i.e. $\lambda = \frac{9 \pm \sqrt{81+24}}{2}$. So we have one strictly positive and one strictly negative eigenvalue and the matrix is not semi-definite.]

11. (a) The Hessian of f is the constant matrix $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, which is positive definite (its determinant is $8+1+1-2-2-2=4$), so f is convex.

(b) Because $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is positive definite, $f(\mathbf{0}) = 0$ and $f(\mathbf{x}) > 0$ for every non-zero $\mathbf{x}$, so that $\mathbf{0}$ is the unique global minimum of f.

(c) Because $f(\mathbf{x}) = (x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_2 + x_3)^2$, f can be of moderate size only if $y_1 = x_1 - x_2$, $y_2 = x_1 - x_3$ and $y_3 = x_2 + x_3$ are all of moderate size. But in this case $x_1 = \frac{1}{2}(y_1 + y_2 + y_3)$ is of moderate size, so x_2 and x_3 also are.

[Alternatively: because $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is positive definite, it must be coercive.]

(d) Because f is coercive, it takes a least value on any closed set, in particular, on the closed set $\{\mathbf{x} : x_1x_2x_3 \leq 1\}$. Because this is not a bounded set, f cannot take a greatest value there.

12. (a) The functions $g_1(\mathbf{x}) = x_1^2+x_2^2+x_3^2-6$ and $g_2(\mathbf{x}) = x_1+x_2+x_3-4$ are continuous, so the constraint set $F = \{\mathbf{x} : g_1(\mathbf{x}) = g_2(\mathbf{x}) = 0\}$ is closed. g_1 is coercive, so F is bounded. On the set $\{\mathbf{x} : g_1(\mathbf{x}) = 0\}$ the function g_2 takes all values between $\pm 3\sqrt{2}$, so F is not empty. The continuous function $f(\mathbf{x}) = x_1x_2x_3$ is therefore bounded and attains its bounds on F, and there will be at least one solution to the problem stated.

Because the functions f, g_1 and g_2 are defined everywhere and are differentiable, all solutions can be found by the Lagrange multiplier method.

(b) The Lagrangean is

$$L(\mathbf{x}, \boldsymbol{\lambda}) = x_1x_2x_3 - \lambda_1(x_1^2+x_2^2+x_3^2-6) - \lambda_2(x_1+x_2+x_3-4),$$

so our equations are

$$\begin{aligned}
0 &= \frac{\partial L}{\partial x_1} = x_2x_3 - 2\lambda_1 x_1 - \lambda_2 \\
0 &= \frac{\partial L}{\partial x_2} = x_1x_3 - 2\lambda_1 x_2 - \lambda_2 \\
0 &= \frac{\partial L}{\partial x_3} = x_1x_2 - 2\lambda_1 x_3 - \lambda_2 \\
0 &= \frac{\partial L}{\partial \lambda_1} = 6 - x_1^2 - x_2^2 - x_3^2 \\
0 &= \frac{\partial L}{\partial \lambda_2} = 4 - x_1 - x_2 - x_3.
\end{aligned}$$

The top three equations imply that

$$x_1x_2x_3 = 2\lambda_1 x_1^2 + \lambda_2 x_1 = 2\lambda_1 x_2^2 + \lambda_2 x_2 = 2\lambda_1 x_3^2 + \lambda_2 x_3,$$

so that x_1, x_2, x_3 all satisfy a quadratic equation $2\lambda_1 x^2 + \lambda_2 x = c$ and cannot all be different.

case 1 If $x_1 = x_2$, then

$$2x_1^2 + x_3^2 = 6, \quad 2x_1 + x_3 = 4,$$

so $2x_1^2 + (4-2x_1)^2 = 6$ and $6x_1^2 - 16x_1 + 16 = 6$ and $3x_1^2 - 8x_1 + 5 = 0$ and $x_1 = 1$ or $\frac{5}{3}$.

case 1a If $x_1 = x_2 = 1$ then $x_3 = 2$ and $f(\mathbf{x}) = 2$.

case 1b If $x_1 = x_2 = \frac{5}{3}$ then $x_3 = \frac{2}{3}$ and $f(\mathbf{x}) = \frac{50}{27}$.

By symmetry, we shall get the same values of f when $x_1 = x_3$ or $x_2 = x_3$, so the maximum value of f is 2, attained at the three points $\begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}2\\1\\1\end{pmatrix}$.

13. (a) With no constraints, we look for points at which

$$0 = \frac{\partial f}{\partial x_1} = -2x_2$$
$$0 = \frac{\partial f}{\partial x_2} = 1 + 2x_2 - 2x_1,$$

so we get

$$x_2 = 0,\ x_1 = \tfrac{1}{2};\ \text{not feasible.}$$

(b) With the constraint $x_1^2 - x_2 = 0$ we have $L = x_2 + x_2^2 - 2x_1x_2 - \lambda(x_1^2 - x_2)$, so must solve

$$0 = \frac{\partial L}{\partial x_1} = -2x_2 - 2\lambda x_1$$
$$0 = \frac{\partial L}{\partial x_2} = 1 + 2x_2 - 2x_1 + \lambda$$
$$0 = \frac{\partial L}{\partial \lambda} = x_2 - x_1^2.$$

Multiplying the second equation by $2x_1$ and adding the top equation, we have

$$0 = -2x_2 + 2x_1 + 4x_1x_2 - 4x_1^2 = 2(x_1 - x_2)(1 - 2x_1).$$

If $x_1 = x_2$, then the equation $x_1^2 = x_2$ gives us either

$$x_1 = 0,\ x_2 = 0,\ f = 0$$

or

$$x_1 = 1,\ x_2 = 1,\ f = 0.$$

If $x_1 = \frac{1}{2}$, then we have

$$x_1 = \frac{1}{2},\ x_2 = \frac{1}{4},\ f = \frac{1}{16}.$$

(c) With the constraint $x_1^2 - 1 = 0$, we have $L = x_2 + x_2^2 - 2x_1x_2 - \lambda(x_1^2 - 1)$, so must solve

$$0 = \frac{\partial L}{\partial x_1} = -2x_2 - 2\lambda x_1$$
$$0 = \frac{\partial L}{\partial x_2} = 1 + 2x_2 - 2x_1$$
$$0 = \frac{\partial L}{\partial \lambda} = 1 - x_1^2.$$

The latter two equations give

$$x_1 = 1,\ x_2 = \frac{1}{2};\ \text{not feasible}$$

and

$$x_1 = -1,\ x_2 = -\frac{3}{2};\ \text{not feasible.}$$

(d) With the two constraints $x_1^2 = 1 = x_2$ we have the points

$$x_1 = x_2 = 1,\ f = 0$$

and

$$x_1 = -1,\ x_2 = 1,\ f = 4.$$

(e) Together we see that the minimum is $f = 0$ at $x_1 = x_2 = 0$ and again at $x_1 = x_2 = 1$.

14.

x	$x + 4\cos x$
0	4
3	−0·9600
1·5	1·7829
2·25	−0·2627
1·875	0·6769
2·0625	0·1740
2·1562	−0·0541
2·1094	0·0577
2·1328	0·0012

15.

x	$x + 4\cos x$
0	4
3	−0·9600
2·4194	−0·5820
1·5254	1·7070
2·1921	−0·1362
2·1428	−0·0225
2·1331	0·0006

16.

x	$x+4\cos x$	$1-4\sin x$
0	4	1
−4	−6·6146	−2·0272
−7·2629	−5·0338	4·3213
−6·0980	−2·1664	0·2635
2·1221	0·0270	−2·4074
2·1333	0·0001	

(It looks to me as though we were a bit lucky here to come back to the same root as last time.)

This requires 11 calculations of f and f' together; the false-position method used 7. So false-position has the advantage (and this is normal).

17.

x	$\arctan x$
−1	−0·7854
5	1·3734
1·1829	0·8690
−5·3930	−1·3875
−1·3496	−0·9331
6·9543	1·4280
1·9321	1·0932
−14·4670	−1·5018
−4·9764	−1·3725
95·7676	1·5604
42·1690	1·5471
−6207·5329	−1·5706
−3059·0795	−1·5705
29819115·5582	1·5708
14906476·6767	1·5708
−698215701621091·8750	−1·5708
−349107835902529·0000	−1·5708

(Clearly my Fortran compiler is surrendering, even on double precision.)

18. (a) $f''(x) = \cosh x + \sin x \geq 1 - 1 = 0$ for every x (because $\cosh x \geq \cosh 0 = 1$ for every x). So f is convex.

Yes, f is coercive, because $\lim_{x\to\pm\infty} \cosh x = \infty$, and the $\sin x$ term is bounded.

(b) According to whether you split the left-hand or right-hand sides when you have equal intervals, you can get either of the tables

x	$\cosh x - \sin x$	x	$\cosh x - \sin x$
0	1	0	1
1	0·70161	1	0·70161
2	2·85290	2	2·85290
0·5	0·64820	1·5	1·35491
0·25	0·78401	0·5	0·64820
0·75	0·61304	0·75	0·61304
0·625	0·61666	0·875	0·64033
0·875	0·64033	0·625	0·61666
0·6875	0·61118	0·8125	0·62263
0·65625	0·61302	0·68750	0·61118
		0·71875	0·61117

corresponding to final intervals $[0{\cdot}65625, 0{\cdot}75]$ or $[0{\cdot}6875, 0{\cdot}75]$. (With random choices you can get minor variations.)

19. $f'(x) = 4x^3 - 18x$, $f''(x) = 12x^2 - 18$ so $x_{n+1} = x_n - \frac{4x_n^3 - 18x_n}{12x_n^2 - 18}$.

(i)

x_n	$f(x_n)$	$f'(x_n)$	$f''(x_n)$
0·94	−7·1716	−13·5977	−7·3968
−0·8983	−6·6115	13·2700	−8·3163
0·6973	−4·1401	−11·1957	−12·1646
−0·2230	−0·4451	3·9699	−17·4032
0·0051	−0·0002	−0·0918	−17·9997
0·0000	0·0000.		

Terminate at 0·0000.

(ii)

x_n	$f(x_n)$	$f'(x_n)$	$f''(x_n)$
0·95	−7·3080	−13·6705	−7·1700
−0·9566	−7·3987	13·7175	−7·0184
0·9979	−7·9702	−13·9872	−6·0511
−1·3136	−12·5529	14·5780	2·7077
−6·6976	1608·5380	−1081·2183	520·2990
−4·6196	263·3465	−311·1794	238·0839
−3·3125	21·6489	−85·7675	113·6753
−2·5580	−16·0737	−20·9105	60·5232
−2·2126	−20·0937	−3·4992	40·7446
−2·1267	−20·2500.		

Terminates at −2·1267.

(iii)

x_n	$f(x_n)$	$f'(x_n)$	$f''(x_n)$
0·96	−7·4451	−13·7411	−6·9408
−1·0198	−8·2776	14·1138	−5·5213
1·5365	−15·6738	−13·1475	10·3295
2·8093	−8·7435	38·1179	76·7055
2·3124	−19·5326	7·8341	46·1638
2·1427	−20·2417	0·7796	37·0915
2·1216	−20·2500.		

Terminates at 2·1216.

20. $\nabla f(\mathbf{x}) = \begin{pmatrix} 4x_1^3 - 2x_1x_2 \\ -x_1^2 + 2x_2 \end{pmatrix}$, $\quad \mathbf{H}f(\mathbf{x}) = \begin{pmatrix} 12x_1^2 - 2x_2 & -2x_1 \\ -2x_1 & 2 \end{pmatrix}$,

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad f(\mathbf{x}) = 1,$$

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) = \begin{pmatrix} 10 & -2 \\ -2 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} 0.625 \\ 0.125 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.1194,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.8203 \\ -0.1406 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 4.4375 & -1.25 \\ -1.25 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(2)} = \begin{pmatrix} \frac{265}{624} \\ \frac{175}{2496} \end{pmatrix} \doteq \begin{pmatrix} 0.4247 \\ 0.0701 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.0248,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.2468 \\ -0.0401 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 2.0240 & -0.8494 \\ -0.8494 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(3)} \doteq \begin{pmatrix} 0.2865 \\ 0.0315 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.0051,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.0760 \\ -0.0191 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 0.9222 & -0.5731 \\ -0.5731 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(4)} \doteq \begin{pmatrix} 0.1934 \\ 0.0144 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.0011,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.0234 \\ -0.0087 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 0.4202 & -0.3869 \\ -0.3869 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(5)} \doteq \begin{pmatrix} 0.1306 \\ 0.0065 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.0002,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.0072 \\ -0.0040 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 0.1915 & -0.2611 \\ -0.2611 & 2 \end{pmatrix},$$

$$\mathbf{x}^{(6)} \doteq \begin{pmatrix} 0.0881 \\ 0.0030 \end{pmatrix}, \quad f(\mathbf{x}) \doteq 0.0000,$$

$$\nabla f(\mathbf{x}) \doteq \begin{pmatrix} 0.0022 \\ -0.0018 \end{pmatrix}, \quad \mathbf{H}f(\mathbf{x}) \doteq \begin{pmatrix} 0.0873 & -0.1763 \\ -0.1763 & 2 \end{pmatrix}.$$

(Note how the method does not bring x_1 to zero quickly, but does get x_2 and f and ∇f close to zero.)

21. (a) $f(\mathbf{x}) = (x_1^2 - \frac{1}{2}x_2)^2 + \frac{3}{4}x_2^2$. So if either $|x_2|$ or $|x_1^2 - \frac{1}{2}x_2|$ is large, then f is large. But if both $|x_2|$ and $|x_1^2 - \frac{1}{2}x_2|$ are small, then x_1^2 and $\|\mathbf{x}\|$ are also small.

$\mathbf{H}f(\mathbf{x}) = \begin{pmatrix} 12x_1^2 - 2x_2 & -2x_1 \\ -2x_1 & 2 \end{pmatrix}$ has determinant $20x_1^2 - 4x_2$, which is sometimes negative; so f cannot be convex.

(b)

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad f = 1, \quad \nabla f(\mathbf{x}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	1
0·5	0·25
1	1
0·25	0·4375
0·75	0·0625
0·625	0·12109
0·875	0·26172
0·6875	0·07349
0·8125	0·11450
0·71875	0·06190

$$\mathbf{x}^{(1)} \simeq \begin{pmatrix} -0{\cdot}43750 \\ 0{\cdot}28125 \end{pmatrix}, \quad f \simeq 0{\cdot}06190, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}08887 \\ 0{\cdot}37109 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·06190
1	0·03377
2	0·24812
0·5	0·01824
0·25	0·03276
0·75	0·01852
0·375	0·02366
0·625	0·01652
0·5625	0·01692
0·68750	0·01705
0·59375	0·01660

$$\mathbf{x}^{(2)} \simeq \begin{pmatrix} -0{\cdot}38196 \\ 0{\cdot}04932 \end{pmatrix}, \quad f \simeq 0{\cdot}01652, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}18522 \\ -0{\cdot}04726 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·01652
1	0·00709
2	0·02067
0·5	0·00622
0·25	0·00954
0·75	0·00567
0·625	0·00565
0·5625	0·00586
0·6875	0·00559
0·65625	0·00560

$$\mathbf{x}^{(3)} \simeq \begin{pmatrix} -0{\cdot}25462 \\ 0{\cdot}08181 \end{pmatrix}, \quad f \simeq 0{\cdot}00559, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}02437 \\ 0{\cdot}09878 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·00559
1	0·00400
2	0·02010
0·5	0·00260
0·25	0·00355
0·75	0·00275
0·375	0·00294
0·625	0·00254
0·5625	0·00253
0·53125	0·00256

$$\mathbf{x}^{(4)} \simeq \begin{pmatrix} -0{\cdot}24091 \\ 0{\cdot}02624 \end{pmatrix}, \quad f \simeq 0{\cdot}00253, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}04328 \\ -0{\cdot}00556 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·00253
4	0·00215
8	0·00434
2	0·00107
1	0·00129
3	0·00146
1·5	0·00108
2·5	0·00121
1·75	0·00106
1·625	0·00106
1·875	0·00106
1·6875	0·00106
1·8125	0·00106
1·71875	0·00106

$$\mathbf{x}^{(5)} \simeq \begin{pmatrix} -0{\cdot}16516 \\ 0{\cdot}03596 \end{pmatrix}, \quad f \simeq 0{\cdot}00106, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}00614 \\ 0{\cdot}04465 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·00106
1	0·00093
2	0·00464
0·5	0·00052
0·25	0·00067
0·75	0·00061
0·375	0·00056
0·625	0·00053
0·4375	0·00053
0·5625	0·00052
0·53125	0·00052

$$\mathbf{x}^{(6)} \simeq \begin{pmatrix} -0{\cdot}16190 \\ 0{\cdot}01224 \end{pmatrix}, \quad f \simeq 0{\cdot}00052, \quad \nabla f\mathbf{x} \simeq \begin{pmatrix} -0{\cdot}01301 \\ -0{\cdot}00172 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·00052
4	0·00028
8	0·00060
2	0·00030
6	0·00040
3	0·00027
2·5	0·00028
3·5	0·00027
2·75	0·00027
3·25	0·00027
3·125	0·00027
3·0625	0·00027
3·1875	0·00027
3·15625	0·00027

$$\mathbf{x}^{(7)} \doteq \begin{pmatrix} -0{\cdot}12084 \\ 0{\cdot}01768 \end{pmatrix}, \quad f \doteq 0{\cdot}00027, \quad \nabla f\mathbf{x} \doteq \begin{pmatrix} -0{\cdot}00278 \\ 0{\cdot}02077 \end{pmatrix}$$

t	$f(\mathbf{x} - t\nabla f\mathbf{x})$
0	0·00027
1	0·00025
2	0·00106
0·5	0·00015
0·25	0·00018
0·75	0·00017
0·375	0·00016
0·625	0·00016
0·4375	0·00016
0·5625	0·00015
0·46875	0·00015

$$\mathbf{x}^{(8)} \doteq \begin{pmatrix} -0{\cdot}11945 \\ 0{\cdot}00730 \end{pmatrix}, \quad f \doteq 0{\cdot}00015, \quad \nabla f\mathbf{x} \doteq \begin{pmatrix} -0{\cdot}00507 \\ 0{\cdot}00033 \end{pmatrix}$$

Note: here the three-point method at each iteration starts with $t_0 = 0$, $t_1 = 2^m$, $t_2 = 2^{m+1}$ where m is chosen as follows: first seek $m \le 0$ such that $f(\mathbf{x} - 2^m \nabla f(\mathbf{x})) < f(\mathbf{x})$, starting with $m = 0$, and if that doesn't work then moving downwards until successful. If $m < 0$, stop there. Otherwise, look at $f(\mathbf{x} - 2^{m+1}\nabla f(\mathbf{x}))$ and if this is less than $f(\mathbf{x})$ raise m until it isn't.

Once again, we find that x_1^2 and x_2 are roughly the same size.

22. $L(\mathbf{x}, \lambda) = x_1^3 - x_2^4 - \lambda(x_1^2 + 2x_2^2 - 4)$, so our equations are

$$\frac{\partial L}{\partial x_1} = 3x_1^2 - 2\lambda x_1 = 0,$$
$$\frac{\partial L}{\partial x_2} = -4x_2^3 - 4\lambda x_2 = 0,$$
$$x_1^2 + 2x_2^2 = 4.$$

So we have

$$\text{either } x_1 = 0 \text{ or } \lambda = \frac{3}{2}x_1,$$

$$\text{either } x_2 = 0 \text{ or } \lambda = -x_2^2.$$

case 1 If $x_1 = 0$ then $x_2 = \pm\sqrt{2}$, $f = -4$.
case 2 If $x_2 = 0$ then $x_1 = \pm 2$, $f = \pm 8$.
case 3 If $\frac{3}{2}x_1 = -x_2^2$ then $x_1^2 - 3x_1 - 4 = 0$, that is, $(x_1 - 4)(x_1 + 1) = 4$ and x_1 is either 4 or -1; but as $x_1^2 \le 4$, the first is impossible, and $x_1 = -1$, $x_2 = \pm\sqrt{\frac{3}{2}}$, $f = -1 - \frac{9}{4} = -\frac{13}{4}$.

The global maximum is at $x_1 = 2$, $x_2 = 0$ and the global minimum at $x_1 = -2$, $x_2 = 0$.

23. We have

$$L(\mathbf{x},\lambda) = x_1x_2 + x_3 - \lambda(x_1^2 + x_2^2 + x_3^2 - c),$$

so that the equations become

$$x_2 - 2\lambda x_1 = 0, \quad x_1 - 2\lambda x_2 = 0, \quad 1 - 2\lambda x_3 = 0, \quad x_1^2 + x_2^2 + x_3^2 = c.$$

From the first two we have

$$x_1^2 = 2\lambda x_1 x_2 = x_2^2$$

so $x_1 = \pm x_2$.

case 1 $x_1 = x_2$. In this case $x_1 - 2\lambda x_1 = 0$ so either $x_1 = 0$ or $\lambda = \frac{1}{2}$.

case 1a $x_1 = 0 = x_2$, $x_3 = \pm\sqrt{c}$, $\lambda = \pm\frac{1}{2\sqrt{c}}$, $f = \pm\sqrt{c}$.

case 1b $\lambda = \frac{1}{2}$, $x_3 = 1$; only possible if $c \geq 1$; in this case $2x_1^2 + 1 = c$, $x_1 = x_2 = \pm\sqrt{\frac{c-1}{2}}$, $f = \frac{c-1}{2} + 1 = \frac{c+1}{2}$.

case 2 $x_1 = -x_2$. In this case $x_1 + 2\lambda x_1 = 0$ so either $x_1 = 0$ or $\lambda = -\frac{1}{2}$.

case 2a $x_1 = 0 = x_2$; already done.

case 2b $\lambda = -\frac{1}{2}$, $x_3 = -1$; only possible if $c \geq 1$; in this case $2x_1^2 + 1 = c$, $x_1 = \pm\sqrt{\frac{c-1}{2}}$, $x_2 = \mp\sqrt{\frac{c-1}{2}}$, $f = -\frac{c-1}{2} - 1 = -\frac{c+1}{2}$.

Thus

if $c \leq 1$, the minimum is at $x_1^* = x_2^* = 0$, $x_3^* = -\sqrt{c} = f^*$;

if $c > 1$, the minima are at $x_1^* = \pm\sqrt{(c-1)/2}$, $x_2^* = -x_1^*$, $x_3^* = -1$, $f^* = -\frac{1}{2}(c+1)$.

For the derivative, we have

$$\frac{df^*}{dc} = -\frac{1}{2\sqrt{c}} \text{ if } 0 < c < 1,$$

$$\frac{df^*}{dc} = -\frac{1}{2} \text{ if } c > 1,$$

$$\frac{df^*}{dc} = -\frac{1}{2} \text{ if } c = 1.$$

In all cases, $\frac{df^*}{dc} = \lambda^*$, as required by the envelope theorem.

24. The equations are

$$1 - 2\lambda_1 x_1 = 0, \quad 1 - 2\lambda_1 x_2 - 2\lambda_2 x_2 = 0, \quad 1 - 2\lambda_2 x_2 = 0.$$

So none of x_1, x_2, x_3 can be zero,

$$\lambda_1 = \frac{1}{2x_1}, \quad \lambda_2 = \frac{1}{2x_2}, \quad 1 - \frac{x_2}{x_1} - \frac{x_2}{x_3};$$

multiplying by x_1x_3, $x_1x_3 - x_2(x_3 + x_1) = 0$, as required.

When $\gamma = \delta$, we know that $x_1^2 = x_3^2$ so $x_1 = \pm x_3$. But as $x_1x_3 \neq 0$ we cannot have $x_1 + x_3 = 0$, so

$$x_1 = x_3, \quad 2x_2x_1 = x_1^2, \quad x_1 = 2x_2,$$

$$5x_2^2 = \gamma, \quad x_2 = \pm\sqrt{\gamma/5}, \quad f^* = f^*(\gamma,\gamma) = 5x_2 = \pm\sqrt{5\gamma}.$$

Since we are looking for the minimum, it is $f^*(\gamma,\gamma) = -\sqrt{5\gamma}$ when $x_1 = x_3 = -2\sqrt{\gamma/5}$, $x_2 = -\sqrt{\gamma/5}$.

Now recall that $\frac{\partial f^*}{\partial\gamma} = \lambda_1^* = \frac{1}{2x_1} = -\frac{1}{4}\sqrt{\frac{5}{\gamma}}$.

[Of course $\frac{d}{d\gamma}(-\sqrt{5\gamma}) = -\frac{1}{2}\sqrt{\frac{5}{\gamma}}$; this is because $\frac{d}{d\gamma}f^*(\gamma,\gamma) = \frac{\partial}{\partial\gamma}f^*(\gamma,\delta) + \frac{\partial}{\partial\delta}f^*(\gamma,\delta)$.]

25. (a) f is coercive and continuous and g is continuous, so $F = \{\mathbf{x} : g(\mathbf{x}) = 0\}$ is closed and there is an $\mathbf{x} \in F$ at which f is minimal.

f does not have a maximal value because (for instance) $\begin{pmatrix} \alpha \\ -\frac{1}{\alpha} \\ 0 \end{pmatrix} \in F$ for every $\alpha \neq 0$, while $f\begin{pmatrix} \alpha \\ -\frac{1}{\alpha} \\ 0 \end{pmatrix} \geq \alpha^2$.

(b)(i) The equations $\dfrac{\partial f}{\partial x_i} - \lambda \dfrac{\partial g}{\partial x_i} = 0$ become

$$2x_1^* - \lambda x_2^* - \lambda x_3^* = 0, \quad 10x_2^* - \lambda x_1^* - \lambda x_3^* = 0, \quad 16x_3^* - \lambda x_1^* - \lambda x_2^* = 0,$$

and convert immediately into the equations given.

(ii) Since we certainly cannot have every x_i^* equal to zero, the equation

$$\begin{pmatrix} 2 & -\lambda & -\lambda \\ -\lambda & 10 & -\lambda \\ -\lambda & -\lambda & 16 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathbf{0}$$

has a non-trivial solution, and the matrix must have zero determinant. But

$$\det \begin{pmatrix} 2 & -\lambda & -\lambda \\ -\lambda & 10 & -\lambda \\ -\lambda & -\lambda & 16 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -2\lambda^3 - 28\lambda^2 + 320 = -2(\lambda + 4)(\lambda^2 + 10\lambda - 40)$$

is zero iff $\lambda = -4$ or $-5 \pm \sqrt{65}$.

(iii) Because $g(\mathbf{x}^*) = 0$, at least one of the products $x_i^* x_j^*$ is less than 0, so that x_i^* and x_j^* have opposite signs.

Suppose, if possible, that $x_1^{*2} < x_2^{*2}$. Set $x_1 = x_2^*$, $x_2 = x_1^*$, $x_3 = x_3^*$; then $g(\mathbf{x}) = g(\mathbf{x}^*) = 0$, while

$$f(\mathbf{x}^*) - f(\mathbf{x}) = x_1^{*2} + 5x_2^{*2} - x_2^{*2} - 5x_1^{*2} = 4(x_2^2 - x_1^{*2}) > 0$$

and f is not minimal at $\mathbf{x}^*$. So this can't happen, and $x_1^{*2} \geq x_2^{*2}$. Similarly, $x_2^{*2} \geq x_3^{*2}$. Since at least two of the x_j^* are non-zero, x_1^* and x_2^* are non-zero.

Since $|x_1^*| \geq |x_2^*| \geq |x_3^*|$, both $x_1^* + x_2^*$ and $x_1^* + x_3^*$ must have the same sign as x_1^*, or be zero. Since x_1^*, x_2^* and x_3^* cannot all be of the same sign, we must have x_1^* on one side of zero and x_2^*, x_3^* on the other. But as $0 \neq x_1^* = \frac{1}{2}\lambda(x_2^* + x_3^*)$, $\lambda < 0$.

(iv) Eliminating x_1 from the equations

$$10x_2^* - \lambda x_1^* - \lambda x_3^* = 0, \quad 16x_3^* - \lambda x_1^* - \lambda x_2^* = 0,$$

we get $(16 + \lambda)x_3^* = (10 + \lambda)x_2^*$. Since x_2^* is non-zero, so is x_3^*; since x_2^* and x_3^* have the same sign, so do $16 + \lambda$ and $10 + \lambda$ (and these are both non-zero). But this means that $\lambda \notin [-16, -10]$.

(c) By (b-iii), $\lambda \neq -5 + \sqrt{65}$; by (b-iv), $\lambda \neq -5 - \sqrt{65}$. So (b-ii) tells us that $\lambda = -4$. Accordingly

$$2x_1^* + 4x_2^* + 4x_3^* = 0, \quad 10x_2^* + 4x_1^* + 4x_3^* = 0, \quad 16x_3^* + 4x_1^* + 4x_2^* = 0,$$

so

$$x_1^* = -6x_3^*, \quad x_2^* = 2x_3^*;$$

since $g(\mathbf{x}^*) = 0$, $x_1^* = \pm\frac{3}{2}$, $x_2^* = \mp\frac{1}{2}$, $x_3^* = \mp\frac{1}{4}$ and $f(\mathbf{x}^*) = 4$.

There are just two possible solutions for $\mathbf{x}^*$.

26. (a) Start: $f(1) \simeq 2{\cdot}6209$, $f(4) \simeq 2{\cdot}0391$.

a_n	b_n	u	$f(u)$	v	$f(v)$
1	4	2·45*	0·1393*	2·55*	0·0598*
2·45	4	3·175*	0·1767*	3·275*	0·3017*
2·45	3·2750	2·8125*	−0·0265*	2·9125*	−0·0091*
2·45	2·9125	2·6312*	0·0135*	2·7312*	−0·0197*
2·6312	2·9125	2·7219*	−0·0177*	2·8219*	−0·0261*
2·7219	2·9125				

Final interval $[2{\cdot}7219, 2{\cdot}9125]$.
$(f(2{\cdot}8172) \simeq -0{\cdot}0263.)$

(b) $f'(x) = 1 - 3\sin x$.
Starting from $f'(1) \simeq -1{\cdot}5244$, $f'(4) \simeq 3{\cdot}2704$,

a_n	b_n	u	$f'(u)$
1	4	2·5*	−0·7954*
2·5	4	3·25*	1·3246*
2·5	3·25	2·875*	0·2097*
2·5	2·875	2·6875*	−0·3159*
2·6875	2·875		

Final interval $[2{\cdot}6875, 2{\cdot}875]$.
$(f(2{\cdot}7812) \simeq -0{\cdot}0261.)$

(c) $x_{k+1} = \frac{x_{k-1}f'(x_k) - x_k f'(x_{k-1})}{f'(x_k) - f'(x_{k-1})}$.

x_k	$f'(x_k)$
1	−1·5244*
4	3·2704*
1·9538*	−1·7827*
2·6757*	−0·3478*
2·8506*	0·1393*
2·8006*	−0·0034*

so $x = 2{\cdot}8006$ will serve.
(The actual minimum is $x^* = \pi - \arcsin\frac{1}{3} \simeq 2{\cdot}8018$.)

27. $\nabla f(\mathbf{x}) = \begin{pmatrix} 2x_1 - x_3 \\ 2x_2 \\ -x_1 + 2x_3 \end{pmatrix}$; $\nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$, so the line of first search is $\begin{pmatrix} 1-2t \\ 1-2t \\ t \end{pmatrix}$, with

$$f(\mathbf{x}^{(0)} - t\nabla f(\mathbf{x}^{(0)})) = (1-2t)^2 - (1-2t)t + (1-2t)^2 + t^2 = 2 - 9t + 11t^2.$$

This is minimal when $t = \frac{9}{22}$, so

$$\mathbf{x}^{(1)} = \frac{1}{22}\begin{pmatrix} 4 \\ 4 \\ 9 \end{pmatrix} \simeq \begin{pmatrix} 0{\cdot}1818 \\ 0{\cdot}1818 \\ 0{\cdot}4091 \end{pmatrix},$$

$$\nabla f(\mathbf{x}^{(1)}) = \frac{1}{22}\begin{pmatrix} -1 \\ 8 \\ 14 \end{pmatrix},$$

$$\begin{aligned} f(\mathbf{x}^{(1)} - t\nabla f(\mathbf{x}^{(1)})) &= \frac{1}{484}((4+t)^2 - (4+t)(9-14t) + (4-8t)^2 + (9-14t)^2) \\ &= \frac{1}{484}(77 - 261t + 275t^2) \end{aligned}$$

which is minimal when $t = \frac{261}{550}$, so that

$$\begin{aligned} \mathbf{x}^{(2)} &= \frac{1}{22}\begin{pmatrix} 4 \\ 4 \\ 9 \end{pmatrix} - \frac{261}{22{\cdot}550}\begin{pmatrix} -1 \\ 8 \\ 14 \end{pmatrix} \\ &= \frac{1}{22{\cdot}550}\begin{pmatrix} 2461 \\ 112 \\ 1296 \end{pmatrix} \simeq \begin{pmatrix} 0{\cdot}2034 \\ 0{\cdot}0093 \\ 0{\cdot}1071 \end{pmatrix}. \end{aligned}$$

28. $f'(x) = -e^{-x} + 2x$, $f''(x) = e^{-x} + 2$.

(a) Starting from $f'(-1) \doteq -4{\cdot}7183$, $f'(1) \doteq 1{\cdot}6321$,

a_n	b_n	u	$f'(u)$
-1	1	0	-1
0	1	0·5	0·3935
0	0·5	0·25	$-0{\cdot}2788$
0·25	0·5	0·375	0·0627
0·25	0·375		

($f(0{\cdot}3125) \doteq 0{\cdot}8293$.)

(b)

x_k	$f'(x_k)$
-1	$-4{\cdot}7183$
1	1·6321
0·4860	0·3569
0·3421	$-0{\cdot}0260$
0·3519	0·0004

so $x = 0{\cdot}3519$ will serve.

(c)

x_n	$f(x_n)$	$f'(x_n)$	$f''(x_n)$
0	1	-1	2
0·5	0·8565	0·3935	2·6487
0·3514	0·8272	$-0{\cdot}0008$	2·4211
0·3518	0·8272	0·0001	2·4216

so $x = 0{\cdot}3518$ will do.

(The minimum is $x^* \doteq 0{\cdot}3517$ to four decimal places.)

29. (a) A *convex program* is a problem of the form

minimise $f(\mathbf{x})$
subject to $\mathbf{x} \in C$
and $g_1(\mathbf{x}), \ldots, g_k(\mathbf{x}) \leq 0$

where $C \subseteq \mathbb{R}^n$ is a convex set and f, $g_1, \ldots, g_k$ are continuous convex functions defined (at least) on C.

In this case the *dual problem* is

maximise $\inf_{\mathbf{x}\in C} L(\mathbf{x}, \boldsymbol{\lambda})$
subject to $\lambda_1, \ldots, \lambda_k \leq 0$,

where

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \ldots - \lambda_k g_k(\mathbf{x})$$

for $\mathbf{x} \in C$, $\lambda \in \mathbb{R}^k$.

(b) We can restate (P) as

(P) minimise $x_1^2 + 2x_2^2 + 3x_3^2$
for $\mathbf{x} \in \mathbb{R}^3$
subject to the constraints $-x_1 - x_2 + 2 \leq 0$, $-x_2 - x_3 + 1 \leq 0$.

The function $f(\mathbf{x}) = x_1^2 + 2x_2^2 + 3x_3^2$ has $\mathbf{H}f = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ which has eigenvalues 2, 4 and 6, so is positive definite, and f is convex. The set $C = \mathbb{R}^3$ is certainly convex. The functions $g_1(\mathbf{x}) = -x_1 - x_2 + 2$, $g_2(\mathbf{x}) = -x_2 - x_3 + 1$ are convex because they are affine functions (or because their Hessians are zero). $\mathbb{R}^3$ is convex and the functions f, g_1, g_2 are all continuous. So (P) is a convex program.

(c) The dual problem is
(D) maximise $\inf_{\mathbf{x}\in\mathbb{R}^3} x_1^2 + 2x_2^2 + 3x_3^2 - \lambda_1(-x_1 - x_2 + 2) - \lambda_2(-x_2 - x_3 + 1)$
subject to $\boldsymbol{\lambda} \leq \mathbf{0}$.
For any fixed $\boldsymbol{\lambda}$, we have to minimise

$$x_1^2 + \lambda_1 x_1 + 2x_2^2 + (\lambda_1 + \lambda_2)x_2 + 3x_3^2 + \lambda_2 x_3 - 2\lambda_1 - \lambda_2.$$

The minimum will be achieved when $x_1 = -\frac{1}{2}\lambda_1$, $x_2 = -\frac{1}{4}(\lambda_1 + \lambda_2)$ and $x_3 = -\frac{1}{6}\lambda_2$, and its value there is

$$\begin{aligned} h(\boldsymbol{\lambda}) &= -\frac{1}{4}\lambda_1^2 - \frac{1}{8}(\lambda_1 + \lambda_2)^2 - \frac{1}{12}\lambda_2^2 - 2\lambda_1 - \lambda_2 \\ &= -\frac{3}{8}\lambda_1^2 - \frac{5}{24}\lambda_2^2 - \frac{1}{4}\lambda_1\lambda_2 - 2\lambda_1 - \lambda_2. \end{aligned}$$

To find the maximum of this (ignoring the constraint $\boldsymbol{\lambda} \leq \mathbf{0}$ for the moment), differentiate:

$$\begin{aligned} \frac{\partial h}{\partial \lambda_1} &= 0 = -\frac{3}{4}\lambda_1 - \frac{1}{4}\lambda_2 - 2 \\ \frac{\partial h}{\partial \lambda_2} &= 0 = -\frac{1}{4}\lambda_1 - \frac{5}{12}\lambda_2 - 1, \end{aligned}$$

which give us $\lambda_1^* = -\frac{7}{3}$, $\lambda_2^* = -1$. (To check that this is indeed a maximum, look at $\mathbf{H}h = -\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{12} \end{pmatrix}$; since $\det \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{12} \end{pmatrix} = \frac{1}{4} > 0$, $\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{5}{12} \end{pmatrix}$ is positive definite, $\mathbf{H}h$ is negative definite and any stationary point of h must be a maximum.) Since the maximum found has $\boldsymbol{\lambda}^* \leq \mathbf{0}$, this is the solution of (D), with

$$\lambda_1^* = -\frac{7}{3},\ \lambda_2^* = -1,\ x_1 = \frac{7}{6},\ x_2 = \frac{5}{6},\ x_3 = \frac{1}{6},$$

and value $\frac{17}{6}$. To check that this is also a solution of (P), note

either that $g_i(\mathbf{x}) = 0$ for every i, so that we have $h(\boldsymbol{\lambda}^*) = L(\mathbf{x}, \boldsymbol{\lambda}^*) = f(\mathbf{x})$ and $(\mathbf{x}, \boldsymbol{\lambda}^*)$ is a dual optimal pair;

or that $\mathbf{x}$ is an interior point of $C = \mathbb{R}^3$, f and all the g_i are differentiable, and the function $\mathbf{x}(\lambda) = \begin{pmatrix} -\frac{1}{2}\lambda_1 \\ -\frac{1}{4}(\lambda_1 + \lambda_2) \\ -\frac{1}{6}\lambda_2 \end{pmatrix}$ is differentiable, so that $(\mathbf{x}(\boldsymbol{\lambda}^*), \boldsymbol{\lambda}^*)$ must be a dual optimal pair;

or that f must have a minimum (because f is coercive and the feasible set is closed) and there is a point $\mathbf{z} \in C$ at which $g_j(\mathbf{z}) < 0$ for both j (e.g., $\mathbf{z} = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$), so that there must be a dual optimal pair, which must be the one we've found.

30. (a) We can express (P) as

$$\text{minimise } (x_1+1)^2 + x_2^2 \text{ for } \mathbf{x} \in C \text{ subject to } -x_1 - 2x_2 + 2 \le 0.$$

As $f(\mathbf{x}) = (x_1+1)^2 + x_2^2$ and $g(\mathbf{x}) = -x_1 - 2x_2 + 2$ are continuous convex functions (their Hessians being $\mathbf{H}f = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{H}g = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, which are both positive semi-definite), and C is a convex set, this is a convex program.

(b) The dual problem is now

$$\text{(D) find } \max_{\lambda \le 0} \inf_{\mathbf{x} \in C} L(\mathbf{x}, \lambda)$$

where $L(\mathbf{x}, \lambda) = (x_1+1)^2 + x_2^2 - \lambda(-x_1 - 2x_2 + 2)$.

(c) We have

$$\begin{aligned} L(\mathbf{x}, \lambda) &= x_1^2 + 2x_1 + 1 + x_2^2 + \lambda x_1 + 2\lambda x_2 - 2\lambda \\ &= (x_1 + 1 + \tfrac{1}{2}\lambda)^2 + (x_2 + \lambda)^2 - \tfrac{5}{4}\lambda^2 - 3\lambda. \end{aligned}$$

(i) If $\lambda \le 0$ and $1 + \frac{1}{2}\lambda \ge 0$, that is, $-2 \le \lambda \le 0$, $h(\lambda) = \min_{\mathbf{x} \in C} L(\mathbf{x}, \lambda)$ is attained when $x_1 = 0$ and $x_2 = -\lambda$, and takes the value $1 - 2\lambda - \lambda^2 = 2 - (\lambda+1)^2$; this is maximal for $\lambda = -1$ and takes the value 2.
(ii) If $\lambda \le -2$ then $h(\lambda) = \min_{\mathbf{x} \in C} L(\mathbf{x}, \lambda)$ is attained when $x_1 = -1 - \frac{1}{2}\lambda$ and $x_2 = -\lambda$, and takes the value $-\frac{5}{4}\lambda^2 - 3\lambda = -\frac{5}{2}(\lambda + \frac{6}{5})^2 + \frac{9}{5}$; this is maximal when $\lambda = -2$ and takes the value 1.
Putting these together, $\max_{\lambda \le 0} \min_{\mathbf{x} \in C} L(\mathbf{x}, \lambda) = 2$, attained when $\lambda = -1$, $x_1 = 0$, $x_2 = 1$.

(d) Our Lagrangean is now

$$L(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = (x_1+1)^2 + x_2^2 - \lambda_1(x_1 - u_1^2) - \lambda_2(x_2 - u_2^2) - \lambda_3(x_1 + 2x_2 - 2 - u_3^2).$$

Equating the partial derivatives to zero, we get

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 0 = 2x_1 + 2 - \lambda_1 - \lambda_3 \\ \frac{\partial L}{\partial x_2} &= 0 = 2x_2 - \lambda_2 - 2\lambda_3 \\ \frac{\partial L}{\partial u_1} &= 0 = 2\lambda_1 u_1 \\ \frac{\partial L}{\partial u_2} &= 0 = 2\lambda_2 u_2 \\ \frac{\partial L}{\partial u_3} &= 0 = 2\lambda_3 u_3. \end{aligned}$$

(i) If $u_1 = 0$ then $x_1 = 0$ so $x_2 \ge 1$ (because $x_1 + 2x_2 \ge 2$) and $u_2 \ne 0$, $\lambda_2 = 0$, $\lambda_3 = x_2 \ne 0$, $u_3 = 0$, $x_1 + 2x_2 = 2$ so $x_1 = 0$, $x_2 = 1$ and $(x_1+1)^2 + x_2^2 = 2$.
(ii) If $u_1 \ne 0$ then $\lambda_1 = 0$, $\lambda_3 = 2x_1 + 2 > 0$, $u_3 = 0$, $x_1 = 2 - 2x_2$.
(α) If $u_2 = 0$ then $x_2 = 0$, $x_1 = 2$, $(x_1+1)^2 + x_2^2 = 9$.
(β) If $u_2 \ne 0$ then $\lambda_2 = 0 = \lambda_1$ so $2x_1 + 2 = x_2$; together with $x_1 = 2 - 2x_2$ we get $x_2 = \frac{6}{5}$, $x_1 = -\frac{2}{5}$ which is infeasible, since $x_1 = u_1^2$.
Thus $\min_{\mathbf{x} \in C, x_1 + 2x_2 \ge 2} (x_1+1)^2 + x_2^2 = 2$, attained when $x_1 = 0$, $x_2 = 1$.

(e) The solution $\lambda = -1$, $\mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ provided by (c) (in fact, by part (i) of (c)) is a feasible point for (P) with $g(\mathbf{x}) = 0$, so at this point $h(\lambda) = f(\mathbf{x})$ and both must be at the optimal values.

31. (a) Express the constraints in the form $1 - x_1 \leq 0$, $x_1^2 + x_2^2 - 1 \leq 0$. $f(\mathbf{x}) = x_2$ and $g_1(\mathbf{x}) = 1 - x_1$ are convex because they are linear, so have zero Hessians; $g_2(\mathbf{x}) = x_1^2 + x_2^2 - 1$ is convex because its Hessian is the positive definite matrix $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$; all these functions are continuous and $\mathbb{R}^2$ is convex.

(b) If $x_1 \geq 1$ and $x_1^2 + x_2^2 \leq 1$ then $x_1 = 1$, $x_2 = 0$, so we have a unique feasible point $\mathbf{x}^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and this must be the optimum, with $f(\mathbf{x}) = 0$.

(c) The dual problem is
(D) maximise $h(\boldsymbol{\lambda}) = \inf_{\mathbf{x} \in \mathbb{R}^2} L(\mathbf{x}, \boldsymbol{\lambda})$
subject to $\boldsymbol{\lambda} \leq \mathbf{0}$,
where

$$L(\mathbf{x}, \boldsymbol{\lambda}) = x_2 - \lambda_1(x_1^2 + x_2^2 - 1) - \lambda_2(1 - x_1).$$

(d) When $\lambda_1 = 0$, $L(\mathbf{x}, \boldsymbol{\lambda}) = x_2 - \lambda_2(1 - x_1)$ is unbounded below (since x_2 can be anything), so $h(\boldsymbol{\lambda})$ has to be interpreted as $-\infty$.
When $\lambda_1 < 0$,

$$L = -\lambda_1(x_1 - \frac{\lambda_2}{2\lambda_1})^2 + \frac{\lambda_2^2}{4\lambda_1} - \lambda_1(x_2 - \frac{1}{2\lambda_1})^2 + \frac{1}{4\lambda_1} + \lambda_1 - \lambda_2$$

which takes its minimal when $x_1 = \dfrac{\lambda_2}{2\lambda_1}$, $x_2 = \dfrac{1}{2\lambda_1}$ with

$$\begin{aligned} h(\boldsymbol{\lambda}) &= \frac{\lambda_2^2}{4\lambda_1} - \lambda_2 + \frac{1}{4\lambda_1} + \lambda_1 \\ &= \frac{1}{4\lambda_1}(1 + (\lambda_2 - 2\lambda_1)^2) \\ &< 0. \end{aligned}$$

If we take $\lambda_2 = 2\lambda_1$ and let $\lambda_1 \to -\infty$, we see that $h(\boldsymbol{\lambda}) \doteq 0$, so that $\sup_{\boldsymbol{\lambda} \leq 0} h(\boldsymbol{\lambda}) = 0$.

32. (a) Setting $f(x) = e^x(2 + \sin x)$, $f'(x) = e^x(2 + \sin x + \cos x)$ and $f''(x) = e^x(2 + 2\cos x) \geq 0$ for every $x \in \mathbb{R}$, so f is convex. Setting $g(x) = e^x$, $g''(x) = e^x \geq 0$ for every $x \in \mathbb{R}$, so g is convex. f and g are both continuous, and $\mathbb{R}$ is convex. Thus (P) is a convex program.

(b) The dual problem is
(D) maximise $\inf_{x \in \mathbb{R}} e^x(2 + \sin x) - \lambda e^x$
for $\lambda \leq 0$.

(c) If $\lambda = 0$, then $L(x, \lambda) = e^x(2 + \sin x)$ is greater than 0 for every x and converges to 0 as $x \to -\infty$, so

$$h(0) = \inf_{x \in \mathbb{R}} e^x(2 + \sin x) = 0,$$

but the infimum is never attained.
If $\lambda < 0$, then $L = e^x(2 + \sin x - \lambda)$, $\dfrac{\partial L}{\partial x} = e^x(2 + \sin x + \cos x - \lambda)$ is never 0. Since $L \geq e^x \geq 0$ for every x, and $\lim_{x \to -\infty} L(x, \lambda) = 0$, we again have $h(\lambda) = 0$.

(d) 0 is not the optimal value for the primal (P), because the condition '$e^x \leq 0$' is never attained; there is no feasible point.

33. (a) C is convex because if $\mathbf{x}$, $\mathbf{y} \in C$, $\alpha \in [0,1]$ and $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ then
either one of x_2, y_2 is greater than 0 and $0 < \alpha < 1$ so $z_2 = \alpha x_2 + (1-\alpha)y_2 > 0$ and $\mathbf{z} \in C$
or $\alpha \in \{0,1\}$ so $\mathbf{z}$ is equal to one of $\mathbf{x}$, $\mathbf{y}$ and belongs to C
or $\mathbf{x} = \mathbf{y} = \mathbf{0}$ so $\mathbf{z} = \mathbf{0} \in C$.

(b) The dual problem is
(D) maximise $h(\lambda) = \inf_{x \in C} x_1 - \lambda x_2$
subject to $\lambda \leq 0$.

(c) The solution to (P) is $\mathbf{x}^* = \mathbf{0}$. But in (D) we have $h(\lambda) = -\infty$ for every λ, because we can take x_1 arbitrarily far to the left.

34. (a) $\mathbf{H}g = \frac{2}{x_1^3}\begin{pmatrix} x_2^2 & -x_1x_2 \\ -x_1x_2 & x_1^2 \end{pmatrix}$ has eigenvalues 0 and $\frac{2(x_1^2+x_2^2)}{x_1^3}$ which are non-negative if $x_1 > 0$, which is true everywhere on C, so g is convex. Since C is also convex, (P) is a convex program.

Since

$$\{\mathbf{x} : \mathbf{x} \in C,\, g(\mathbf{x}) \leq 0\} = \{\mathbf{x} : x_2 = 0,\, x_1 \geq 1\},$$

the optimum for (P) is $\mathbf{x}^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $f(\mathbf{x}^*) = 1$.

(b) The dual problem is
(D) maximise $\inf_{\mathbf{x} \in C} x_1 - \lambda\frac{x_2^2}{x_1}$
for $\lambda \leq 0$.

Now if $\lambda \leq 0$, $\mathbf{x} \in C$ then $x_1 - \lambda\frac{x_2^2}{x_1} \geq 0$, so $h(\lambda) \geq 0$. If $\lambda = 0$, then $h(0) = \inf_{\mathbf{x} \in C} x_1 = 0$. If $\lambda < 0$, then, given any $\epsilon > 0$, we can set

$$x_1 = \frac{\epsilon}{2}, \quad x_2 = \frac{\epsilon}{2\sqrt{-\lambda}}$$

to get a point $\mathbf{x} \in C$ such that $L(\mathbf{x}, \lambda) = \epsilon$; so $h(\lambda) = \inf_{\mathbf{x} \in C} L(\mathbf{x}, \lambda)$ is again 0.

35. (a) C_0 is a right circular cone with apex $\mathbf{0}$ and axis $x_1 = x_2 = 0$. So it is convex. (It is the case that the function $\mathbf{x} \mapsto \sqrt{x_1^2 + x_2^2} - x_3$ is convex, so that the set C_0 where this function is at most zero must be convex.)

(b) $x_2^2 \leq x_3^2$ so $-x_2 \leq |x_2| \leq |x_3| = x_3$ and $0 \leq x_2 + x_3$.

(c) is elementary.

(d) $(\sqrt{\alpha^2+1} - \alpha)(\sqrt{\alpha^2+1} + \alpha) = 1$ so $\lim_{\alpha \to \infty} \sqrt{\alpha^2+1} - 1 = \lim_{\alpha \to \infty} \frac{1}{\sqrt{\alpha^2+1}+1} = 0$.

36. (a) Defining C_0 as in question 35, $C = C_0 \cap \{\mathbf{x} : x_1 \geq -1\}$ is the intersection of convex sets, so is convex. The functions $\mathbf{x} \mapsto x_1$ and $\mathbf{x} \mapsto x_2 + x_3$ are linear, therefore continuous and convex, so (P) is a convex program.

(b) If $\mathbf{x} \in C$, then $x_2 + x_3 \geq 0$, and if $x_2 + x_3 = 0$ then $x_2 = -x_3$, $x_1 = 0$. Thus the optimal value for (P) is 0, being the value of f at any feasible point.

(c) The dual problem is
(D) maximise $h(\lambda) = \inf_{\mathbf{x} \in C} x_1 - \lambda(x_2 + x_3)$ for $\lambda \leq 0$.

(d) If $\lambda = 0$ then $h(0) = \inf_{\mathbf{x} \in C} x_1 = -1$ because (by the definition of C) $x_1 \geq -1$ for every $\mathbf{x} \in C$, while (for instance) $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \in C$.

If $\lambda < 0$ then $x_1 - \lambda(x_2 + x_3) \geq x_1 \geq -1$ for every $\mathbf{x} \in C$, since $x_2 + x_3 \geq 0$ for every $\mathbf{x} \in C$. Thus $h(\lambda) \geq -1$. On the other hand, for any $\alpha \geq 0$, $\begin{pmatrix} -1 \\ -\alpha \\ \sqrt{\alpha^2 + 1} \end{pmatrix} \in C$, so that

$$h(\lambda) \leq -1 - \lambda(-\alpha + \sqrt{\alpha^2 + 1}) \to -1$$

as $\alpha \to \infty$, and $h(\lambda) = -1$.