Nonlinear programming: unconstrained extrema of functions of several variables.

1. $f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 3x_1 - 6x_2$.
   Gradient $\nabla f = (2x_1 + x_2 - 3,\ x_1 + 2x_2 - 6)$.
   Critical point $P = (0, 3)$.
   Hessian $D^2 f = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, positive definite.
   So, P is a strict local (and absolute) minimizer, where $f = -9$.

2. $f(x_1, x_2) = x_1^3 + 3x_1x_2^2 - 15x_1 - 12x_2$.
   Gradient $\nabla f = (3x_1^2 + 3x_2^2 - 15,\ 6x_1x_2 - 12)$.
   Critical points: intersection of a circle $x_1^2 + x_2^2 = 5$ and a hyperbola $x_1x_2 = 2$, i.e. $P_1 = (1, 2)$, $P_2 = (2, 1)$; $P_3 = (-1, -2)$, $P_4 = (-2, -1)$.
   Hessian $D^2 f = 6 \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}$. At the points P_1, P_3, P_2, P_4 respectively it is:
   $6 \begin{bmatrix} \pm 1 & \pm 2 \\ \pm 2 & \pm 1 \end{bmatrix}$: saddle points; $6 \begin{bmatrix} \pm 2 & \pm 1 \\ \pm 1 & \pm 2 \end{bmatrix}$: local min and max, with $f = \mp 28$, respectively.
   There is no global min. or max.: setting $x_2 = 0$ makes f go to $\pm\infty$.

3. $f(x_1, x_2) = (2x_1^2 + x_2^2)e^{-(x_1^2 + x_2^2)}$.
   Gradient $\nabla f = \left(2x_1(2 - 2x_1^2 - x_2^2)e^{-(x_1^2+x_2^2)},\ 2x_2(1 - 2x_1^2 - x_2^2)e^{-(x_1^2+x_2^2)}\right)$.
   Critical points: $P_0 = (0, 0)$, $P_1 = (0, 1)$, $P_2 = (0, -1)$; $P_3 = (1, 0)$, $P_4 = (-1, 0)$ (note that simultaneously one cannot have $2 - 2x_1^2 - x_2^2 = 0$ and $1 - 2x_1^2 - x_2^2 = 0$).
   Hessian $D^2 f = e^{-(x^2+y^2)} \begin{bmatrix} 4 - 20x^2 - 2y^2 + 4x^2(2x^2 + y^2) & 4xy(2x^2 + y^2 - 12) \\ 4xy(2x^2 + y^2 - 3) & 2 - 4x^2 - 10y^2 + 4y^2(2x^2 + y^2) \end{bmatrix}$.
   Thus, it's easy to check, as the non-diagonal entries vanish at each critical point, P_0 is a local minimum, with $f = 0$; P_3 and P_4 are local maxima with $f = 2e^{-1}$, the rest are saddles.

   As certainly we have $f \geq 0$, the origin is the absolute minimum. As f is continuous and away from the origin, due to the presence of the vanishing exponential, $f \to 0$, $P_{3,4}$ yield absolute maxima.

4. $f(x_1, x_2, x_3) = x_1 + \dfrac{x_2}{x_1} + \dfrac{x_3}{x_2} + \dfrac{2}{x_3}$.
   Gradient $\nabla f = (1 - x_2/x_1^2,\ 1/x_1 - x_3/x_2^2,\ 1/x_2 - 2/x_3^2)$.
   Critical points: $P_1 = (-\sqrt[4]{2}, \sqrt{2}, -\sqrt[4]{8})$, $P_2 = (\sqrt[4]{2}, \sqrt{2}, \sqrt[4]{8})$.
   Hessian $D^2 f = \begin{bmatrix} 2x_2/x_1^3 & -1/x_1^2 & 0 \\ -1/x_1^2 & 2x_3/x_2^3 & -1/x_2^2 \\ 0 & -1/x_2^2 & 4/x_3^3 \end{bmatrix}$.
   Evaluating the Hessian at the points $P_1$and P_2 and applying the Sylvestre criterion, one verifies that the Hessian is negative definite at P_1 and positive definite at P_2, the former thus being a local maximizer and the latter a local minimizer with the values of $f = \mp 4\sqrt[4]{2}$, respectively. Note that the local maximum is actually smaller than the local minimum, as they occur on different branches of f, separated by the zero values of x_1, x_2, x_3.

   In addition, it is easy to see that f can be made arbitrary large positive/negative, so absolute min. or max. don't exist.

5. $f(x_1, x_2) = 2 - \sqrt[3]{x_1^2 + x_2^2}$. No analysis is necessary to see that $x_1 = x_2 = 0$ is the absolute maximum, where $f = 2$.

6. $f(x_1, x_2, x_3) = x_1 x_2^2 x_3^3 (1 - x_1 - 2x_2 - 3x_3),\ \boldsymbol{x} > 0$.
$\nabla f = \left(x_2^2 x_3^3 (1 - 2x_1 - 2x_2 - 3x_3),\ 2x_2 x_1 x_3^3 (1 - x_1 - 3x_2 - 3x_3),\ 3x_1 x_2^2 x_3^2 (1 - x_1 - 2x_2 - 4x_3) \right)$.
The only critical point with $\boldsymbol{x} > 0$ is $x_1 = x_2 = x_3 = 1/7$. Evaluating the Hessian at this points shows that it is negative definite there, so this point is a maximizer, with $f(1/7, 1/7, 1/7) = 7^{-7}$.

   This is clearly the absolute maximum, in the domain $\boldsymbol{x} > 0$.

7. $2x^2 + 2y^2 + z^2 + 8yz - z + 8 = 0$, for an implicit function $z(x, y)$.
Implicit-differentiate with respect to x and with respect to y:
$4x + 2zz_x + 8yz_x - z_x = 0$ and $4y + 2zz_y + 8z + 8yz_y - z_y = 0$.
Critical points correspond to $\nabla z(x, y) = (z_x, z_y) = (0, 0)$, therefore $x = 0$ and $y = -2z$. Plugging it into the equation yields $z = 1, -8/7$, so $y = -2, 16/7$, respectively.

   To characterise these critical points, differentiate the above expressions once again with respect to x and y and get: $4 + 2z_x z_x + 2zz_{xx} + 8yz_{xx} - z_{xx} = 0$, $2z_y z_x + 2zz_{xy} + 8z_x + 8yz_{xy} - z_{xy} = 0$, $4 + 2z_y z_y + 2zz_{yy} + 8z_y + 8z_y + 8yz_{yy} - z_{yy} = 0$. Now plug in $x = 0$ and $y = -2, 16/7$, knowing then in this case $z_x = z_y = 0$ (these are critical points!) and $z = 1, -8/7$ respectively. Get $z_{xx} = z_{yy} \pm 4/15$ while $z_{xy} = 0$. Thus, the critical point $(x, y) = (0, -2)$ is a local minimizer, yielding $z = 1$, the critical point $(x, y) = (0, 16/7)$ is a local maximizer, yielding $z = -8/7$. The minimizer and the maximizer correspond to a pair of different branches of the surface $z = z(x, y)$, defined implicitly.

Convex functions

1. What convexity properties do the following functions have:
   (a) $f(x) = x^2 - 10x + 2,\ x \in \mathbb{R}$. Convex: $f'' = 2 > 0$.
   (b) $f(x) = \ln x,\ x > 0$. Concave: $f'' = -1/x^2 < 0$.
   (c) $f(x) = e^x,\ x \in \mathbb{R}$. Convex: $f'' = e^x > 0$.
   (d) $f(x_1, x_2) = x_1^2 + 3x_2^2 - x_1 x_2,\ \boldsymbol{x} \in \mathbb{R}^2$. Convex: the Hessian is positive definite.
   (e) $f(x_1, x_2, x_3) = -x_1^2 - x_2^2 - 2x_3^2 + \dfrac{1}{2} x_1 x_2,\ \boldsymbol{x} \in \mathbb{R}^3$. Concave: the Hessian is negative definite.

2. It suffices to show that $f(x, y) = e^{x^2 + 2y^2}$ is a convex function: we know that any is sublevel set $\{(x, y) : f(x, y) \le C\}$ is convex for any C. Now, $f_{xx} = (2 + 4x^2) e^{x^2 + 2y^2} > 0$, $f_{xy} = 8xy e^{x^2 + 2y^2}$, $f_{yy} = (4 + 16y^2) e^{x^2 + 2y^2}$. So the Hessian, by the Sylvester criterion, is easily to be positive definite, and f is convex.

3. True-false questions:
   (a) False: a linear function is both convex and concave.
   (b) True: by definition, if $f(\theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2) \le \theta_1 f(\boldsymbol{x}_1) + \theta_2 f(\boldsymbol{x}_2)$ and $g(\theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2) \le \theta_1 g(\boldsymbol{x}_1) + \theta_2 g(\boldsymbol{x}_2)$, adding these two inequalities yields $[f + g](\theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2) \le \theta_1 [f + g](\boldsymbol{x}_1) + \theta_2 [f + g](\boldsymbol{x}_2)$
   (c) False in general: take $f = x^2$ and $g = x - 1$ in $\mathbb{R}^1$ for a counterexample.

4. Inequalities problems.
   (a) By convexity of $f(x) = x^p$ for $p > 1$ ($f''(x) = p(p - 1) x^{p-2} > 0$ for $x > 0$) and definition of convex function. (with $\theta_1 = \theta_2 = \frac{1}{2}$.)
   (b) By concavity of $f(x) = x^{1/p}$ for $p > 1$ (as $f''(x) < 0$ for $x > 0$, using the generalized definition of convexity with $\theta_1 = \ldots = \theta_3 = 1/3$).

(c) Geometric mean is less than the arithmetic mean bug greater than the harmonic one (Cauchy inequality) with four terms.

(d) Same, with five terms, using the formulae for the sum of the geometric and arithmetic progressions.

(e) By Cauchy-Shwartz inequality: $\sum_{i=1}^n x_i = \sum_{i=1}^n x_i \cdot 1 \leq \sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n 1^2} = \sqrt{n \sum_{i=1}^n x_i^2}.$

(f) Same with the Hölder inequality:

$$\sum_{i=1}^n x_i = \sum_{i=1}^n 1 \cdot x_i \leq \left(\sum_{i=1}^n 1^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^n x_i^q\right)^{\frac{1}{q}} = n^{\frac{1}{p}} \left(\sum_{i=1}^n x_i^q\right)^{\frac{1}{q}}.$$

(g) Follows from generalised Jensen's inequality $f\left(\frac{\sum p_i x_i}{\sum p_i}\right) \geq \frac{\sum p_i f(x_i)}{\sum p_i}$ for a concave f. Take $f = \ln x$, $p_i = \Delta s_i$, $x_i = g(s_i)$ for a partition $s_1, s_2, \ldots, s_n$ of $[0,1]$ (all $s_i \in (0,1)$, $s_{i+1} > s_i$), so $\sum p_i = 1$. Then get

$$\ln\left(\sum g(s_i)\Delta s_i\right) \geq \sum \ln[g(s_i)]\Delta s_i,$$

and the integral inequality in question as $n \to \infty$.

5. Let $x \in A+A$ and $n(x)$ the number of representations of x as $x = a_1 + a_2$. Given x, the number of ordered quadruples which are (a_1, a_2, a_3, a_4) solutions of the equation $x = a_1 + a_2 = a_3 + a_4$ equals $n^2(x)$, as independently $x = a_1 + a_2$ has $n(x)$ choices and $x = a_3 + a_4$ has $n(x)$ choices. So, the total number of solutions of the equation $a_1 + a_2 = a_3 + a_4$ is

$$\sum_{x \in A+A} n^2(x)$$

By Cauchy-Schwartz (the version in (e) above) we have

$$|A+A| \sum_{x \in A+A} n^2(x) \geq \left(\sum_{x \in A+A} n(x)\right)^2$$

But in the right-hand side $\sum_{x \in A+A} n(x) = N^2$, since this is just the total number of *all* ordered pairs (a_1, a_2) from A – each such pair gives *some* sum $x = a_1 + a_2$, and then summation is taken over all possible x.

So, in the problem's notation we have

$$XE \geq N^4,$$

this does it.

6. Introduce the characteristic function $S(x,y,z)$ of the set S, which equals 1 if the point $(x,y,z) \in S$ and $S(x,y,z) = 0$ otherwise. Let $S_1(x,y)$, $S_2(y,z)$, $S_3(z,x)$ be characteristic functions of the projections of the set S onto the xy, yz, zx-planes, respectively. Then

$$S(x,y,z) \leq S_1(x,y) S_2(y,z) S_3(z,x).$$

Indeed, $S(x,y,z) = 1$ only if $S_1(x,y), S_2(y,z), S_3(z,x)$ are all equal to 1. Besides, $\sum_{x,y,z} S(x,y,z) = N$. Use this and Cauchy-Scwartz applied twice:

First, we apply it to summation in (x, y):

$$N \leq \sum_{x,y} S_1(x,y)\left(\sum_z S_2(y,z)S_3(z,x)\right) \leq \left(\sum_{x,y} S_1^2(x,y)\right)^{1/2} \cdot \left(\sum_{x,y}\left(\sum_z S_2(y,z)S_3(z,x)\right)^2\right)^{1/2}.$$

In the first multiplier,

$$\sum_{x,y} S_1^2(x,y) = \sum_{x,y} S_1(x,y) = |P_{xy}(S)|,$$

where $|P_{xy}(S)|$ denotes the size of the projection of S onto the xy-plane.

In the second multiplier, apply Cauchy-Scwartz to the summation in z:

$$\left(\sum_z S_2(y,z)S_3(z,x)\right)^2 \leq \sum_z S_2^2(y,z) \cdot \sum_z S_3^2(z,x) = \sum_z S_2(y,z) \cdot \sum_z S_3(z,x)$$

So, we have

$$\sum_{x,y}\left(\sum_z S_2(y,z)S_3(z,x)\right)^2 \leq \sum_{x,y}\sum_z S_2(y,z)\cdot\sum_z S_3(z,x) = \sum_{y,z} S_2(y,z)\cdot\sum_{x,z} S_3(z,x) = |P_{yz}(S)||P_{xz}(S)|,$$

where $|P_{yz}(S)|$, $|P_{xz}(S)|$ denote the size of the projection of S onto the yz and xz-planes respectively. Thus, altogether

$$N^2 \leq |P_{xy}(S)||P_{uz}(S)||P_{xz}(S)|,$$

the product of the sizes of the three projections, hence one of them must be is greater than $N^{2/3}$.

Note, the inequality is sharp, take S as the "lattice cube" $[1, \ldots, M] \times [1, \ldots, M] \times [1, \ldots, M]$. The size of each projection is M^2, while S itself has size M^3.