Solutions to Problem Sheet 2

(2.2) For this problem we generalize the notion of convexity to function not necessarily defined on all of $\mathbb{R}^n$. Denote by $\mathrm{dom} f$ the *domain* of f, i.e., the set of $\boldsymbol{x}$ on which $f(\boldsymbol{x})$ attains a finite value. A function f is called *convex*, if $\mathrm{dom} f$ is a convex set and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom} f$ and $\lambda \in [0,1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}).$$

Which of the following functions are convex?

(a) $f(x) = \log(x)$ on $\mathbb{R}_{++}$ (the positive real numbers); **(2)**

(b) $f(\boldsymbol{x}) = x_1 x_2$ on $\mathbb{R}^2_{++}$; **(2)**

(c) $f(\boldsymbol{x}) = x_1/x_2$ on $\mathbb{R}^2_{++}$; **(2)**

(d) $f(\boldsymbol{x}) = \max_i x_i$ on $\mathbb{R}^n$. **(2)**

Solution.

(a) Not convex. Take $f'(x) = 1/x$ and $f''(x) = -1/x^2$;

(b) Not convex. The Hessian is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and this is not positive definite. Note that to check whether this matrix is positive definite, one has to very the definition $x^T A x > 0$ for all x, not only x in the domain. Any other justification will do.

(c) Not convex. The Hessian is $\begin{pmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{pmatrix}$ is not positive semidefinite.

(d) Convex. For $\boldsymbol{x} \neq \boldsymbol{y}$ we have

$$\max_i (\lambda x_i + (1-\lambda) y_i) \leq \max_i \lambda x_i + \max_j (1-\lambda) y_j = \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}).$$

(2.3) Determine the order of convergence of each of the following sequences (if they converge at all).

$$\text{(a)}\ \ x_k = \frac{1}{k!}, \quad \text{(b)}\ \ x_k = 1 + (0.3)^{2^k}, \quad \text{(c)}\ \ x_k = 2^{-k}, \quad \text{(d)}\ \ x_k = 1/k$$

(4)

Solution.

(a) The sequence converges to 0. We have

$$x_{k+1} = \frac{1}{(k+1)!} = \frac{1}{k+1} x_k,$$

so that $\lim_{k\to\infty} \frac{x_{k+1}}{x_k} = 0$. The convergence is *superlinear*.

(b) The sequence converges to 1. We have

$$|x_{k+1} - 1| = (0.3)^{2^{k+1}} = \left((0.3)^{2^k}\right)^2 = |x_k - 1|^2,$$

so that the convergence is quadratic.

(c) The sequence converges to 0. Moreover,

$$x_{k+1} = \frac{1}{2^{k+1}} = \frac{1}{2}x_k,$$

so that the sequence converges linearly.

(d) The sequence converges to 0. We have the identity

$$x_{k+1} = \frac{1}{k+1} = \frac{k}{k+1}\frac{1}{k} = \frac{k}{k+1}x_k,$$

which means that for any fixed constant $c < 1$ there is a k such that $1 > k/(k+1) > c$, and therefore $x_{k+1} > cx_k$. It follows that the sequence does not converge linearly (or to any higher order).

(2.4) Consider the function on $\mathbb{R}^2$, $f(\boldsymbol{x}) = (x_1^2 + x_2)^2$. Show that the direction $\boldsymbol{p} = (1, -1)^\top$ is a descent direction at $\boldsymbol{x}_0 = (0, 1)^\top$, and determine a step length α that minimizes $f(\boldsymbol{x}_0 + \alpha\boldsymbol{p})$. **(3)**

Solution. The gradient is

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4(x_1^2 + x_2)x_1 \\ 2(x_1^2 + x_2) \end{pmatrix}.$$

At $\boldsymbol{x}_0 = (0, 1)^\top$, $\nabla f(0, 1) = (0, 2)^\top$. The direction $\boldsymbol{p}$ is a descent direction, if $\langle \nabla f(\boldsymbol{x}_0), \boldsymbol{p}\rangle < 0$. In our case, $\langle \nabla f(\boldsymbol{x}_0), \boldsymbol{p}\rangle = -2 < 0$. The optimal step length along $\boldsymbol{p}$ is the minimizer of

$$f(\boldsymbol{x}_0 + \alpha\boldsymbol{p}) = (\alpha^2 + 1 - \alpha)^2.$$

The minimum is the minimum of the quadratic function $\alpha^2 + 1 - \alpha$. Computing the derivative and setting it to zero, $2\alpha - 1 = 0$, we get the optimal step length $\alpha = 1/2$.

(2.5) Consider an optimisation problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t.} \quad x \in \Omega \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex on Ω and Ω is a convex set. Show that the optimal solution is unique (assuming it exists). **(5)**

Solution. Assume that the statement is not true and that there are two optimal points $\boldsymbol{x}$ and $\boldsymbol{y}$ in Ω, such that $f(\boldsymbol{x}) = f(\boldsymbol{y}) = p$. Since Ω is convex, for all $\lambda \in [0, 1]$, $\lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in \Omega$. Since f is strictly convex, for $\lambda \in (0, 1)$ we have

$$f(\lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) < \lambda f(\boldsymbol{x}) + (1 - \lambda)f(\boldsymbol{y}) = p,$$

in contradiction to the assumption that $\boldsymbol{x}$ and $\boldsymbol{y}$ are minimizers.

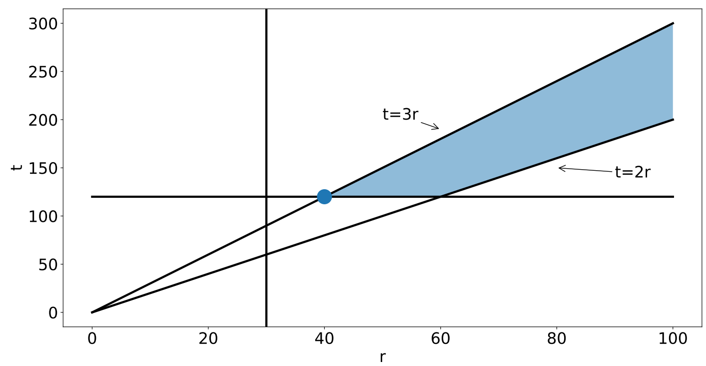

Figure 1: Linear Programming

(2.6) A local politician is budgeting for her media campaign. She will distribute her funds between TV ads and radio ads. She has been given the following advice by her campaign advisers:

- She should run at least 120 TV ads and at least 30 radio ads.
- The number of TV ads she runs should be at least twice the number of radio ads she runs but not more than three times the number of radio ads she runs.

The cost of a TV ad is $\pounds 8000$ and the cost of a radio ad is $\pounds 2000$.

Formulate the respective linear program and sketch the feasible set. What do you observe for the solution.**(5)**

Solution. Let t be the number of TV ads and r the number of radio ads. Then $t \geq 120$, $r \geq 30$ and $2r \leq t \leq 3r$. The objective function is $f(t, r) = 8000t + 2000r$, which she wishes to minimize. The minimizer is plotted on the image as the point with the smallest (r, t) coordinate in the feasible (coloured) region.