Unconstrained optimisation

To apply the SONC and SOSC to optimisation, we first need to have an understanding of quadratic forms. Questions 1 and 2 will familiarise you with quadratic forms. Following this, in Question 3, you will see that the conditions for local minimisers can be applied quite reasonably when the feasible region is not constrained.

1. Consider the following matrices:

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} -2 & 2 \\ 2 & -5 \end{pmatrix}, \quad \mathbf{A}_4 = \begin{pmatrix} -1 & -1 \\ -1 & 3 \end{pmatrix}.$$

   (a) By explicitly computing the matrix $\mathbf{x}^T\mathbf{A}_i\mathbf{x}$, write out the quadratic form corresponding to each $\mathbf{A}_i$.

   (b) Determine the definiteness of each matrix. That is, identify whether the matrix is positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite.

   (c) Determine the definiteness of $\mathbf{A}_4$ on the subspace $\{\mathbf{x} : x_1 + x_2 = 0\}$.

2. Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

   (a) Write out the quadratic form corresponding to $\mathbf{A}$.

   (b) Determine the definiteness of $\mathbf{A}$.

   (c) Determine the definiteness of $\mathbf{A}$ on the subspace $\{\mathbf{x} : x_1 + x_2 + x_3 = 0\}$.

3. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$, be given by

$$f(\mathbf{x}) = \tfrac{1}{3}(x_1^3 + x_2^3) - 4(x_1 + 4x_2).$$

   (a) Determine ∇f and $D^2 f$.

   (b) Suppose $\mathbf{x}^*$ is a minimiser of f. Explain why the FONC implies $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

   (c) Find all points $\mathbf{p} \in \mathbb{R}^2$ that satisfy the FONC for f.

   (d) For each point $\mathbf{p}$ found in (c), determine the definiteness of $D^2 f(\mathbf{p})$, and make suitable deductions about these points using the SONC and the SOSC. Are any of these points global extremisers?

Constrained optimisation

When the function is constrained to a region Ω, it takes more effort to identify candidates for local minimisers, as the feasible directions at a point on the boundary must also taken into consideration. In general, it can be quite difficult to directly apply the conditions for local minimisers to a constrained optimisation problem. Throughout this subject, we will look at some particular cases where general methods can be applied. For now, the following questions illustrate the main underlying principles.

4. For each of the following, sketch the set Ω and determine the feasible directions at the point $\mathbf{x}^* = \begin{pmatrix} 1 & 2 \end{pmatrix}^T$. Then, determine whether $\mathbf{x}^*$ is (i) a local minimiser (ii) not a local minimiser or (iii) possibly a local minimiser, of a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ with the given properties, subject to $\mathbf{x} \in \Omega$.

   (a) $\Omega = \{\mathbf{x} : x_1 \geqslant 1\}$ and $\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} 1 & 1 \end{pmatrix}$.

   (b) $\Omega = \{\mathbf{x} : x_1 \geqslant 1, x_2 \geqslant 2\}$ and $\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} 1 & 0 \end{pmatrix}$.

   (c) $\Omega = \{\mathbf{x} : x_1 \geqslant 1, x_2 \geqslant 2\}$ and $\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} 1 & 1 \end{pmatrix}$.

   (d) $\Omega = \{\mathbf{x} : x_1 \geqslant 0, x_2 \geqslant 0\}$, $\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} 0 & 0 \end{pmatrix}$, and $D^2 f(\mathbf{x}^*) = \mathbf{I}_2$ (where $\mathbf{I}_2$ is the 2×2 identity matrix).

   (e) $\Omega = \{\mathbf{x} : x_1 \geqslant 1, x_2 \geqslant 2\}$, $\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} 1 & 0 \end{pmatrix}$, and $D^2 f(\mathbf{x}^*) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

5. Consider $f\colon \mathbb{R}^2 \to \mathbb{R}$, given by $f(\mathbf{x}) = -x_2^2$, subject to $|x_2| \leqslant x_1^2$ and $x_1 \geqslant 0$.
   (a) Sketch the feasible region Ω.
   (b) Does $\mathbf{0}$ satisfy the FONC?
   (c) Is $\mathbf{0}$ a strict local maximiser?
6. Consider $f\colon \mathbb{R}^2 \to \mathbb{R}$, given by $f(\mathbf{x}) = 5x_2$, subject to $\mathbf{x} \in \Omega = \{\mathbf{x} : x_1^2 + x_2 \geqslant 1\}$.
   (a) Sketch the feasible region Ω.
   (b) Does $\mathbf{x}^* = \begin{pmatrix}0 & 1\end{pmatrix}^T$ satisfy the FONC?
   (c) Does $\mathbf{x}^* = \begin{pmatrix}0 & 1\end{pmatrix}^T$ satisfy the SONC?
   (d) Is $\mathbf{x}^* = \begin{pmatrix}0 & 1\end{pmatrix}^T$ a local minimiser?
7. Consider $f\colon \mathbb{R}^2 \to \mathbb{R}$, given by $f(\mathbf{x}) = -3x_1$, subject to $\mathbf{x} \in \Omega = \{\mathbf{x} : x_1 + x_2^2 \leqslant 2\}$.
   (a) Sketch the feasible region Ω.
   (b) Does $\mathbf{x}^* = \begin{pmatrix}2 & 0\end{pmatrix}^T$ satisfy the FONC?
   (c) Does $\mathbf{x}^* = \begin{pmatrix}2 & 0\end{pmatrix}^T$ satisfy the SONC?
   (d) Is $\mathbf{x}^* = \begin{pmatrix}2 & 0\end{pmatrix}^T$ a local minimiser?

Additional exercises

8. If a function $f\colon \mathbb{R}^n \to \mathbb{R}$ can be written as $f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{h}(\mathbf{x})$, then the product rule tells us that
$$Df(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T D\mathbf{h}(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T D\mathbf{g}(\mathbf{x}).$$
   Use the product rule to show each of the following facts.
   (a) If $\mathbf{A}$ is an arbitrary $1 \times n$ matrix and $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, then $Df(\mathbf{x}) = \mathbf{A}$.
   (b) If $\mathbf{A}$ is an arbitrary $n \times 1$ matrix and $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}$, then $Df(\mathbf{x}) = \mathbf{A}^T$.
   (c) If $\mathbf{A}$ is an arbitrary $n \times n$ matrix and $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, then $Df(\mathbf{x}) = \mathbf{x}^T(\mathbf{A} + \mathbf{A}^T)$.

   For each of the above, what is special about the case $n = 1$?
9. Why are some of the functions in Question 8 written using bold (e.g. $\mathbf{g}$), but some are written in italics (e.g. f)?
10. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by
$$f(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix}1 & 2\\ 4 & 7\end{pmatrix} \mathbf{x} + \mathbf{x}^T \begin{pmatrix}3\\ 5\end{pmatrix} + 6.$$
    (a) Determine ∇f and $D^2 f$. *Use the results of Question 8.*
    (b) Suppose $\mathbf{x}^*$ is a minimiser of f. Explain why the FONC implies $\nabla f(\mathbf{x}^*) = \mathbf{0}$.
    (c) Find all points $\mathbf{p} \in \mathbb{R}^2$ that satisfy the FONC for f.
    (d) Do the points in (c) also satisfy the SONC?
    (e) Make a conclusion on the basis of your previous answers.
11. Find and classify all extrema of the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by
$$f(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix} \mathbf{x} + \mathbf{x}^T \begin{pmatrix}3\\ 2\end{pmatrix} + 1.$$
12. Given n real numbers $x_1, \ldots, x_n$, find the number $\bar{x}$ such that the sum of the squared difference between $\bar{x}$ and the numbers is minimised.

13. Consider the following optimisation problem:

$$\begin{aligned}\text{maximise} \quad & z = c_1x_1 + c_2x_2 \\ \text{subject to} \quad & x_1 + x_2 \leqslant 1 \\ & \mathbf{x} \geqslant \mathbf{0},\end{aligned}$$

for some constants $c_1 > c_2 > 0$. Use the FONC to show that the maximiser $\mathbf{x}^*$

(a) cannot lie in the interior of the constraint set

(b) cannot lie on one of the line segments $L_1 = \{\mathbf{x} : x_1 = 0 \leqslant x_2 < 1\}$, $L_2 = \{\mathbf{x} : x_2 = 0 \leqslant x_1 < 1\}$, or $L_3 = \{\mathbf{x} : 0 \leqslant x_1 < 1, x_2 = 1 - x_1\}$.

(c) is possibly given by $\begin{pmatrix}1 & 0\end{pmatrix}^T$,

and show that $\begin{pmatrix}1 & 0\end{pmatrix}^T$ is a global maximum.

More MATLAB

14. Consider the following optimisation problem:

$$\begin{aligned}\text{maximise} \quad & f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ \text{subject to} \quad & |x_2 - 1| \leqslant 2 \\ & |x_1| \leqslant 2.\end{aligned}$$

The function $f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ is known as the Rosenbrock function and is a commonly used function for testing optimisation methods.

(a) Implement a function `rosenbrock` in MATLAB that computes the Rosenbrock function.

(b) Determine $\nabla f(\mathbf{x})$. Then implement a function `grosenbrock` in MATLAB with two outputs, `g1` and `g2`, which are the first and second components of $\nabla f(\mathbf{x})$, respectively.

(c) Using the MATLAB function `surf`, create a surface plot in the region $\mathcal{R} = \{\mathbf{x} : |x_1| \leqslant 2, |x_2 - 1| \leqslant 2\}$.
See Question 5 of Workshop 1.

(d) Using the MATLAB function `contour`, plot the c-level sets of f, for $c = 0.7,\ 7,\ 70,\ 200,\ 700$ (in one diagram).
Read the documentation for the `contour` *function by entering* `help contour` *in the command window.*

(e) *Read the documentation for the* `quiver` *function by entering* `help quiver` *in the command window.*

(f) Observe that there is a turning point on the 0.7-level set. Construct a plot to inspect some contours and the gradient of f near this turning point.

(g) Produce a plot which shows the Rosenbrock-function has a minimum.