1.3 First-order conditions for local minimisers

The first-order conditions in this section and the second-order conditions in the next section are based on the description given by Chong and Żak in *An Introduction to Optimization* [3].

A function $f\colon \mathbb{R} \to \mathbb{R}$ can be approximated by polynomial functions near points where the function is continuous and differentiable (sufficiently many times). For a point a meeting those conditions, Taylor's theorem tells us that, near a,

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2} + \dots.$$

Without going into too much detail about limits, differentiability, the meaning of $\approx$, et cetera, in this section, we will provide the formulas for the first and second order Taylor series for vector valued functions of many variables.

Background: derivatives in higher dimensions

Consider a differentiable vector-valued function $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$. We let f_i denote the i-th coordinate of $\mathbf{f}$. Then, the **derivative** of $\mathbf{f}$ at $\mathbf{a}$, also called the **Jacobian**, is given by

$$D\mathbf{f}(\mathbf{a}) = \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{pmatrix}(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}(\mathbf{a}).$$

Note that if f is a real-valued function (i.e., if $m = 1$), then $Df(\mathbf{a})$ is a $1 \times n$ row matrix. For a real-valued function $f\colon \mathbb{R}^n \to \mathbb{R}$, the **gradient** of f (at $\mathbf{a}$) is the column vector

$$\nabla f(\mathbf{a}) = Df(\mathbf{a})^T = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}(\mathbf{a})$$

For notational convenience, sometimes we will write $\nabla^T f(\mathbf{a})$ instead of $\nabla f(\mathbf{a})^T$ and $D^T f(\mathbf{a})$ instead of $Df(\mathbf{a})^T$.

Example: derivatives in higher dimensions

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

$$f(x,y,z) = x^2 + 3y^2 z + \sin(xy)$$

Then

$$Df(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x + y\cos(xy) & 6yz + x\cos(xy) & 3y^2 \end{pmatrix}$$

and so

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2x + y\cos(xy) \\ 6yz + x\cos(xy) \\ 3y^2 \end{pmatrix}$$

A vector function $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if it preserves vector addition and scalar multiplication; i.e., if $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ and $f(c\mathbf{v}) = cf(\mathbf{v})$. A function $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ is **affine** if it can be written as $\mathbf{f}(\mathbf{x}) = \mathbf{l}(\mathbf{x}) + \mathbf{b}$, where $\mathbf{l}$ is linear and $\mathbf{b} \in \mathbb{R}^m$.

Theorem 1 (First order Taylor)
If $\mathbf{f}\colon \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\mathbf{a}$, then the best affine approximation to $\mathbf{f}$ near $\mathbf{a}$ is given by

$$\mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}) \tag{1.4}$$

In the above theorem, the $\mathcal{D}$ written in script font denotes the domain of the function, while the D typeset in italics denotes the Jacobian. Note that this construction involves matrix multiplication: $D\mathbf{f}(\mathbf{a})$ is an $m \times n$ matrix and $(\mathbf{x} - \mathbf{a})$ is an $n \times 1$ column matrix. Consequently, $D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$ is an $m \times 1$ column matrix.

Background: directional derivatives

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a differentiable real-valued function. We now consider the change of the function f in a given direction $\mathbf{d} \in \mathbb{R}^n$.

Starting with the first order Taylor series (1.4), replacing simultaneously $\mathbf{x} \to \mathbf{x} + \alpha\mathbf{d}$ and $\mathbf{a} \to \mathbf{x}$ we get the formula

$$f(\mathbf{x} + \alpha\mathbf{d}) = f(\mathbf{x}) + \alpha Df(\mathbf{x})\mathbf{d},$$

which is valid for small α. This actually means that

$$\lim_{\alpha \to 0} \frac{f(\mathbf{x} + \alpha\mathbf{d}) - f(\mathbf{x})}{\alpha} = Df(\mathbf{x})\mathbf{d}. \tag{1.5}$$

The left hand side of this expression (1.5) is called the **directional derivative** of f in the direction $\mathbf{d}$, denoted by $\frac{\partial f}{\partial \mathbf{d}}(\mathbf{x})$, and the right-hand side of the expression gives us a formula for calculating it. When $\|\mathbf{d}\| = 1$, the directional derivative $Df(\mathbf{x})\mathbf{d}$ is the **rate of change** of f at $\mathbf{x}$ in the direction of $\mathbf{d}$.

Using the chain rule, the right hand side of (1.5) can be shown to equal $\frac{d}{d\alpha} f(\mathbf{x} + \alpha\mathbf{d}) \mid_{\alpha=0}$.

Example: directional derivatives

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

$$f(x, y, z) = x^2 + 3y^2 z + \sin(xy).$$

From the previous example, we have

$$Df(\mathbf{x}) = \begin{pmatrix} 2x + y\cos(xy) & 6yz + x\cos(xy) & 3y^2 \end{pmatrix}.$$

For a given direction $\mathbf{d} = (d_1, d_2, d_3)^T$, the directional derivative of f in the direction $\mathbf{d}$ is

$$\begin{aligned} Df(\mathbf{x})\mathbf{d} &= \begin{pmatrix} 2x + y\cos(xy) & 6yz + x\cos(xy) & 3y^2 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \\ &= (2x + y\cos(xy)) \times d_1 + (6yz + x\cos(xy)) \times d_2 + (3y^2) \times d_3 \\ &= d_1\,(2x + y\cos(xy)) + d_2\,(6yz + x\cos(xy)) + 3d_3 y^2. \end{aligned}$$

In particular, at $\mathbf{x} = (0, 1, 2)$ in the direction $\mathbf{d} = (1, 1, -1)$, we have

$$Df(\mathbf{x})\mathbf{d} = 1 \times (0 + \cos(0)) + 1 \times (6 \times 2 + 0\cos(0)) + 3 \times -1 \times 1^2 = 10$$

For a corresponding rate of change, we need to normalise $\mathbf{d}$ so that it has length 1. Taking the unit vector $\hat{\mathbf{d}} = \frac{1}{\|\mathbf{d}\|}\mathbf{d} = \frac{1}{\sqrt{3}}(1, 1, -1)$, the rate of change at $\mathbf{x} = (0, 1, 2)$ in the direction $\hat{\mathbf{d}}$ is given by

$$Df(\mathbf{x})\hat{\mathbf{d}} = \frac{1}{\sqrt{3}} Df(\mathbf{x})\mathbf{d} = \frac{10}{\sqrt{3}}.$$

Recall the general form of the optimisation problem we consider:

$$\begin{aligned} \text{minimise} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Omega. \end{aligned} \tag{1.6}$$

We assume that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. A nonzero vector $\mathbf{d} \in \mathbb{R}^n$ is a **feasible direction** at $\mathbf{x} \in \Omega$ if there exists $\beta > 0$ such that $\mathbf{x} + \alpha\mathbf{d} \in \Omega$ for all $0 \leqslant \alpha \leqslant \beta$. The following theorem gives a First Order Necessary Condition for $\mathbf{x}^*$ to be a local minimiser.

Theorem 2 (FONC)
If $\mathbf{x}^*$ is a local minimiser for (1.6), then for all feasible directions $\mathbf{d}$ at $\mathbf{x}^*$ we have

$$\nabla f(\mathbf{x}^*)^T \mathbf{d} \geqslant 0.$$

By symmetry, if $\mathbf{x}^*$ is a local maximiser for (1.6), then for all feasible directions $\mathbf{d}$ at $\mathbf{x}^*$ we have

$$\nabla f(\mathbf{x}^*)^T \mathbf{d} \leqslant 0.$$

Proof. Since $\mathbf{d}$ is a feasible direction at $\mathbf{x}^*$ there exist $\beta > 0$ such that $\mathbf{x}^* + \alpha\mathbf{d} \in \Omega$ for all $0 < \alpha < \beta$. If $\mathbf{x}^*$ is a local minimiser, we must have $f(\mathbf{x}^* + \alpha\mathbf{d}) \geqslant f(\mathbf{x}^*)$. Then, dividing by $\alpha > 0$, taking a limit, and using (1.5) yields

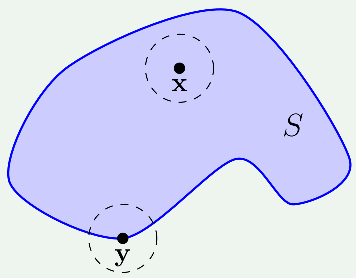
$$0 \leqslant \lim_{\alpha \to 0} \frac{f(\mathbf{x}^* + \alpha\mathbf{d}) - f(\mathbf{x}^*)}{\alpha} = D\mathbf{f}(\mathbf{x})\mathbf{d} = \nabla f(\mathbf{x})^T \mathbf{d}.$$

□

Background: Interior points

Let S be a subset of $\mathbb{R}^n$. A point $\mathbf{x} \in S$ is an **interior point** of S if there exists some $\varepsilon > 0$ such that $\mathcal{B}_\varepsilon(\mathbf{x}) \subset S$. In other words, $\mathbf{x}$ is an interior point if S contains a neighbourhood of $\mathbf{x}$. The set of all interior points of S is called the **interior** of S.

A point $\mathbf{x}$ (not necessarily in S) is a **boundary point** of S if every neighbourhood of $\mathbf{x}$ contains at least one point in S and one point not in S. In the diagram below, the point $\mathbf{x}$ depicts an interior point of S and the point $\mathbf{y}$ depicts a boundary point of S.



A set S is **open** if all points in S are interior points, and a set is **closed** if it contains its boundary.

The union of a set with its boundary is called the **closure** of the set.

In $\mathbb{R}$, open intervals are examples of open sets, and closed intervals are examples of closed sets.

Corollary 3

If $\mathbf{x}^*$ is a local extremiser in the interior of Ω, then $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Proof. Every vector $\mathbf{d}$ is feasible at a point in the interior. By Theorem 2, if $\mathbf{x}^*$ is a local minimiser, then for all $\mathbf{d}$, we have both $\nabla f(\mathbf{x})^T\mathbf{d} \geqslant 0$ and $\nabla f(\mathbf{x})^T(-\mathbf{d}) \geqslant 0$, and hence $\nabla f(\mathbf{x})^T\mathbf{d} = 0$. A similar argument applies if $\mathbf{x}^*$ is a local maximiser. □

Example: applying the FONC

Since the FONC is a *necessary* condition, it allows us to make conclusions of the following form:

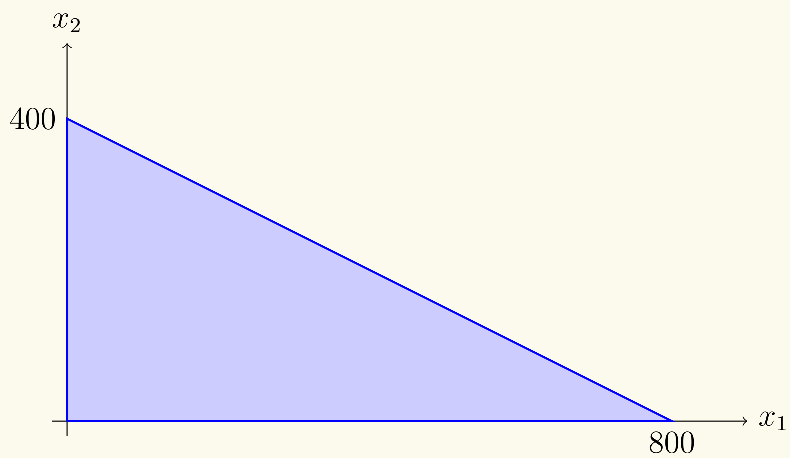
- $\mathbf{x}^*$ *might* be a local minimiser, or
- $\mathbf{x}^*$ *is definitely not* a local minimiser.

It does not allow us to make a definitive conclusion about whether a point *is* a local minimiser, although in some cases it can be used to identify possible candidates.

Consider the fencing problem from Section 1.1, rephrased as a minimisation problem:

$$\begin{aligned} \text{minimise} \quad & f(x_1, x_2) = -x_1x_2 \\ \text{subject to} \quad & x_1 + 2x_2 \leqslant 800 \\ & x_1, x_2 \geqslant 0. \end{aligned}$$

Sketching the feasible set for this problem gives the following:



Consider Corollary 3: if a point in the interior is a local minimiser, then we must have $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Here, we have

$$\nabla^T f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -x_2 & -x_1 \end{pmatrix}.$$

So $\nabla f(\mathbf{x}) = \mathbf{0}$ if and only if $x_1 = x_2 = 0$, but the point $(0, 0)$ is not in the interior of Ω. This tells us that there are no local minimisers in the interior of the feasible set, and we deduce that a local minimiser must be on the boundary of the region.

Let's consider whether the top left corner, $\mathbf{x}^* = \begin{pmatrix} 0 & 400 \end{pmatrix}^T$, might a local minimiser. The feasible directions at this point are those $\mathbf{d} = \begin{pmatrix} d_1 & d_2 \end{pmatrix}^T$ such that $d_1 \geqslant 0$ and $d_2 \leqslant -\frac{d_1}{2}$. To see why this is the case, consider the slope of the diagonal of the triangle.

At $\mathbf{x}^* = \begin{pmatrix} 0 & 400 \end{pmatrix}^T$, we have

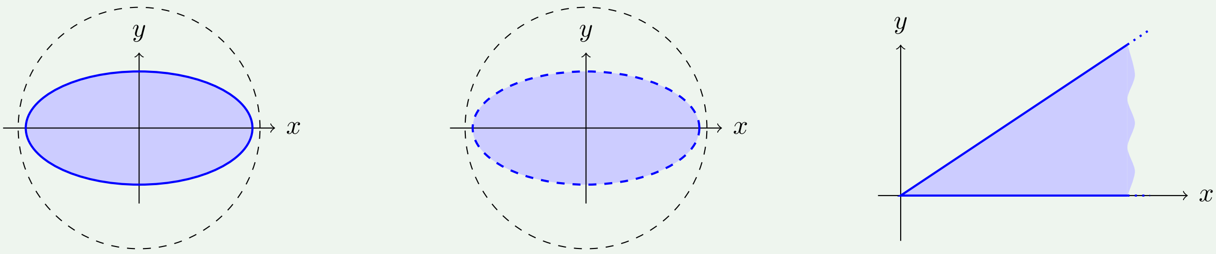
$$\nabla^T f(\mathbf{x}^*) = \begin{pmatrix} -400 & 0 \end{pmatrix} \implies \nabla^T f(\mathbf{x}^*)\mathbf{d} = \begin{pmatrix} -400 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = -400 d_1.$$

But then taking, say, the feasible direction $\mathbf{d} = \begin{pmatrix} 2 & -1 \end{pmatrix}^T$, would give $\nabla^T f(\mathbf{x}^*)\mathbf{d} = -800 < 0$, so the FONC is not satisfied. Thus, $\mathbf{x}^* = \begin{pmatrix} 0 & 400 \end{pmatrix}^T$ is not a local minimiser.

Similar reasoning can be used to exclude all points on the boundary except for the point $\mathbf{x}^* = \begin{pmatrix} 400 & 200 \end{pmatrix}^T$. At $\mathbf{x}^*$, the set of feasible directions can be written as $\{\alpha(2,-1) - \beta(1,2) : \alpha, \beta \in \mathbb{R}, \beta \geqslant 0\}$. This alone would not be sufficient to prove that it is a local minimiser, although it can then be deduced as the only possible outcome from Theorem 4 below.

Background: compact sets

A set is **bounded** if it is contained in a ball of finite radius. A set is **compact** if it both closed and bounded.



The elliptical region given by $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1$ is both closed and bounded, hence compact.

The elliptical region given by $\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1$ is bounded but not closed.

The infinite region given by $0 \leqslant y \leqslant ax$ is closed but not bounded.

Weierstrass's theorem is a result that establishes the existence of local extrema over compact sets, although it does not give us any information about how to find them.

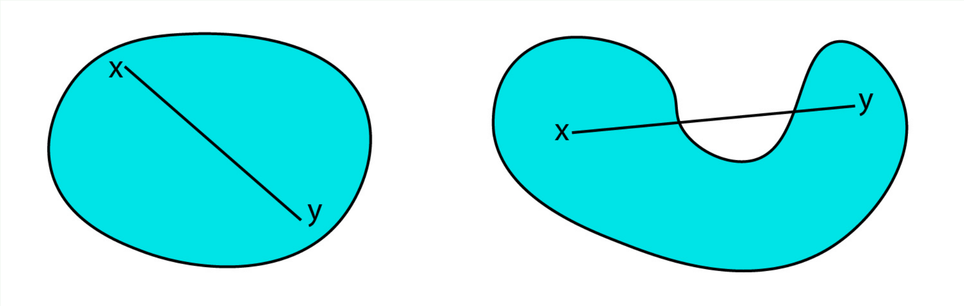
Theorem 4 (Weierstrass)
Let $f\colon \mathcal{D} \to \mathbb{R}$ be a continuous function on a compact set $\mathcal{D}$. Then there exists a local minimiser and a local maximiser of f over $\mathcal{D}$.

Background: convex sets

In $\mathbb{R}^n$ the **line segment** between two points $\mathbf{x}$ and $\mathbf{y}$ is the set

$$\mathcal{S} = \{\alpha\mathbf{x} + (1-\alpha)\mathbf{y} : \alpha \in [0,1]\}.$$

A set $\mathcal{U} \subset \mathbb{R}^n$ is **convex** if, for all $\mathbf{x}, \mathbf{y} \in \mathcal{U}$, the line segment between $\mathbf{x}$ and $\mathbf{y}$ is a subset of $\mathcal{U}$.



In the above Figure the set on the left is convex, the set on the right is not.

The following theorem, not given in [3], gives a First Order Sufficient Condition for $\mathbf{x}^*$ to be a local minimizer of a differentiable function f.

Theorem 5 (FOSC)
Let Ω be convex and let $\mathcal{D}$ be the set of all feasible directions of length 1, at $\mathbf{x}^* \in \Omega$. If $\mathcal{D}$ is closed and if $\mathbf{x}^*$ is a point such that, for all $\mathbf{d} \in \mathcal{D}$,

$$\nabla f(\mathbf{x}^*)^T \mathbf{d} > 0,$$

then $\mathbf{x}^*$ is a strict local minimizer for the optimisation problem (1.6).

By symmetry, if $\nabla f(\mathbf{x}^*)^T \mathbf{d} < 0$ instead, then $\mathbf{x}^*$ is a strict local maximiser.

Proof. We shall prove that there is an ϵ such that, for all $\mathbf{x} \in \mathcal{B}_\epsilon(\mathbf{x}^*) \cap \Omega$, we have $f(\mathbf{x}) \geqslant f(\mathbf{x}^*)$. Let $\mathbf{x} \in \mathcal{B}_\epsilon(\mathbf{x}^*) \cap \Omega$. Then, since Ω is convex, we can write $\mathbf{x} = \mathbf{x}^* + \alpha \mathbf{d}$ with $\alpha \geqslant 0$, and $\mathbf{d}$ a feasible direction of length 1. As f is differentiable, using the Taylor expansion we have

$$\frac{f(\mathbf{x}^* + \alpha \mathbf{d}) - f(\mathbf{x}^*)}{\alpha} = \nabla f(\mathbf{x}^*)^T \mathbf{d} + \alpha(\cdots).$$

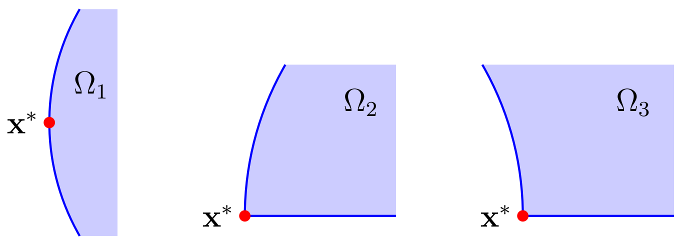
For all feasible $\mathbf{d}$ the first term is greater than 0, so, for small enough $0 < \alpha < \beta(\mathbf{d})$, the quantity on the left is greater than zero. If we take $\epsilon = \beta$, where

$$\beta = \min_{\{\mathbf{d} \text{ feasible}: \|\mathbf{d}\|=1\}} \beta(\mathbf{d}),$$

which exists according to Weierstrass, Theorem 4, then we conclude that $f(\mathbf{x}) > f(\mathbf{x}^*)$, i.e. $\mathbf{x}^*$ is a strict local minimizer. □

Alternatively, in the statement of Theorem 1, we can assume that Ω is **locally convex**, meaning that, for each $\mathbf{x} \in \Omega$, there is a λ such that for all $0 < \delta < \lambda$ the set $\mathcal{B}_\delta(\mathbf{x}^*) \cap \Omega$ is convex. At the end of the proof one should then take $\epsilon = \min(\beta, \lambda)$.

Note that the conditions in the FOSC imply that $\mathbf{x}^*$ is a boundary point, in fact, a corner point. Can you see why? Also, what could go wrong if Ω is not convex? Or if the set of feasible directions of unit length is not closed? Note that for the situations sketched below, only in the last one is the set of feasible directions of length 1 closed.



As the set in the picture on the right is not convex, the theorem 1 is of limited use. It might be the case that the following is true. However, we don't have a proof.

Conjecture 1 (FOSC)
For any Ω (not necessarily convex), let $\mathcal{D}$ be the closure of all feasible directions of length 1, at $\mathbf{x}^* \in \Omega$. If $\mathbf{x}^*$ is a point such that, for all $\mathbf{d} \in \mathcal{D}$,

$$\nabla f(\mathbf{x}^*)^T \mathbf{d} > 0,$$

then $\mathbf{x}^*$ is a strict local minimizer for the optimisation problem (1.6).

By symmetry, if $\nabla f(\mathbf{x}^*)^T \mathbf{d} < 0$ instead, then $\mathbf{x}^*$ is a strict local maximiser.

The next section will describe second-order conditions for local minimisers, based on second derivatives. This section, and the next, serve primarily as a theoretical basis for optimisation. If the objective function is nonlinear and depends on many variables, it is not likely that we will be able to solve the conditions presented in this section and the next. In later parts, we will turn to iterative search methods to find approximate solutions instead.

1.4 Second-order conditions for local minimisers

Recall the general form of the optimisation problem we consider:

$$\begin{aligned} \text{minimise} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \Omega. \end{aligned} \tag{1.7}$$

The second-order conditions for local minimisers are based on taking second derivatives.

Background: second derivatives in higher dimensions

Consider a differentiable real-valued function $f\colon \mathbb{R}^n \to \mathbb{R}$. Recall that the gradient of f (at $\mathbf{a}$)is the column vector

$$\nabla^T f(\mathbf{a}) = Df(\mathbf{a}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} (\mathbf{a}).$$

Furthermore, if ∇f is also differentiable, then the derivative of ∇f is

$$D^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

and is called the **Hessian**.

Example: second derivatives in higher dimensions

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

$$f(x, y, z) = x^2 + 3y^2 z + \sin(xy).$$

From a previous example, we have

$$Df(\mathbf{x}) = \begin{pmatrix} 2x + y\cos(xy) & 6yz + x\cos(xy) & 3y^2 \end{pmatrix}.$$

The Hessian is then

$$D^2 f(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y} \\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}.$$

Note that the first column (or row) is obtained by differentiating each component of $Df(\mathbf{x})$ with respect to x, the second column is obtained by differentiating each coordinate of $Df(\mathbf{x})$ with respect to y, and the third column is obtained by differentiating each coordinate of $Df(\mathbf{x})$ with respect to z.

This gives

$$D^2 f = \begin{pmatrix} 2 - y^2\sin(xy) & \cos(xy) - xy\sin(xy) & 0 \\ \cos(xy) - xy\sin(xy) & 6z - x^2\sin(xy) & 6y \\ 0 & 6y & 0 \end{pmatrix}.$$

Observe that $D^2 f$ is a symmetric matrix, i.e., it is equal to its own transpose.

Theorem 6 (Second order Taylor)
If $f\colon \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}$ is twice differentiable at $\mathbf{a}$, then the best quadratic approximation to f near $\mathbf{a}$ is given by

$$f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \frac{1}{2}(\mathbf{x}-\mathbf{a})^T D^2 f(\mathbf{a})(\mathbf{x}-\mathbf{a}). \tag{1.8}$$

As in the first order case, this construction involves matrix multiplication. Since $D^2f(\mathbf{a})$ is $n \times n$ and $(\mathbf{x}-\mathbf{a})$ is $n \times 1$, applying $D^2f(\mathbf{a})(\mathbf{x}-\mathbf{a})$ yields an $n \times 1$ matrix. Then, since $(\mathbf{x}-\mathbf{a})^T$ is a $1 \times n$ column matrix, $(\mathbf{x}-\mathbf{a})^T D^2 f(\mathbf{a})(\mathbf{x}-\mathbf{a})$ is a 1×1 matrix, which we treat as a scalar.

Next, we give a Second Order Necessary Condition, assuming that f is differentiable twice.

Theorem 7 (SONC)
Suppose that $\mathbf{d}$ is a feasible direction at $\mathbf{x}^*$ such that $\nabla^T f(\mathbf{x}^*)\mathbf{d} = 0$. If $\mathbf{x}^*$ is a local minimiser for the optimisation problem (1.7), then

$$\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d} \geqslant 0.$$

By symmetry, if $\mathbf{x}^*$ is a local maximser, then $\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d} \leqslant 0$ instead.

Proof. Similar to the proof of Theorem 2. Since $\mathbf{d}$ is a feasible direction at $\mathbf{x}^*$ there exist $\alpha > 0$ such that $\mathbf{x}^* + \alpha\mathbf{d} \in \Omega$ for all $0 < \alpha < \alpha^*$. If $\mathbf{x}^*$ is a local minimiser, we must have $f(\mathbf{x}^* + \alpha\mathbf{d}) \geqslant f(\mathbf{x}^*)$. According to the second order Taylor approximation (1.8), replacing simultaneously $\mathbf{x} \to \mathbf{x} + \alpha\mathbf{d}$ and $\mathbf{a} \to \mathbf{x}$, we have

$$f(\mathbf{x}^* + \alpha\mathbf{d}) = f(\mathbf{x}^*) + \alpha Df(\mathbf{x}^*)\mathbf{d} + \frac{\alpha^2}{2}\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d},$$

which is valid in the limit where $\alpha \to 0$. So, as $Df(\mathbf{x}^*)\mathbf{d} = 0$, we obtain

$$0 \leqslant \lim_{\alpha \to 0} \frac{f(\mathbf{x}^* + \alpha\mathbf{d}) - f(\mathbf{x}^*)}{\alpha^2} = \frac{1}{2}\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d}.$$

□

Background: quadratic forms

An $n \times n$ matrix $\mathbf{M}$ is **symmetric** if $M = M^T$. A real **quadratic form** is a function $Q\colon \mathbb{R}^n \to \mathbb{R}$ of the form

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{Q}\mathbf{x},$$

for some symmetric $n \times n$ matrix Q.

A quadratic form Q (or a symmetric matrix $\mathbf{Q}$) is:

- **positive definite** if $Q(\mathbf{u}) > 0$ for all $\mathbf{u} \neq 0$,
- **positive semi-definite** if $Q(\mathbf{u}) \geqslant 0$ for all $\mathbf{u}$,
- **negative definite** if $Q(\mathbf{u}) < 0$ for all $\mathbf{u} \neq 0$,
- **negative semi-definite** if $Q(\mathbf{u}) \leqslant 0$ for all $\mathbf{u}$.

One way that an $n \times n$ symmetric matrix $\mathbf{Q}$ can be checked for positive/negative definiteness is using **Sylvester's criterion** ([3, Thm 3.6]). Let $\mathbf{Q}^{(k)}$ denote the $k \times k$ submatrix taken from the top left corner of Q. These matrices are called the **leading principal minors** of $\mathbf{Q}$. Then let $\Delta_k = \det(\mathbf{Q}^{(k)})$. Sylvester's criterion is that:

- Q is positive definite if and only if $\Delta_k > 0$ for all k.
- Q is negative definite if and only if $(-1)^k \Delta_k > 0$ for all k.

The condition that the leading principal minors are non-negative is a necessary but not a sufficient condition for positive semi-definiteness.

It can be shown that, if the second partial derivatives of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ are all continuous, then the Hessian $D^2 f$ is symmetric. This leads us to the following results.

Corollary 8

If $\mathbf{x}^*$ is a local minimiser in the interior of Ω, and $\nabla^T f(\mathbf{x}^*) = \mathbf{0}$, then $D^2 f(\mathbf{x}^*)$ is positive semi-definite.

If $\mathbf{x}^*$ is a local maximiser in the interior of Ω, and $\nabla^T f(\mathbf{x}^*) = \mathbf{0}$, then $D^2 f(\mathbf{x}^*)$ is negative semi-definite.

Proof. Similar to the proof of Corollary 3. □

A slight modification gives a sufficient condition.

Theorem 9 (SOSC)

Let either $\mathbf{x}^*$ be an interior point of Ω, or let Ω be locally convex at $\mathbf{x}^*$. If

- $\nabla^T f(\mathbf{x}^*) = \mathbf{0}$ and
- $D^2 f(\mathbf{x}^*)$ is positive definite, or, when $\mathbf{x}^*$ is a boundary point, $\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d} > 0$ for all $\mathbf{d}$ in the set of feasible directions of length 1, which should be closed.

then $\mathbf{x}^*$ is a local minimiser for the optimisation problem (1.7).

By symmetry, replacing positive definite with negative definite, and $>$ with $<$, in the statement above will result in a local maximiser.

Proof. If $\mathbf{x}^*$ is interior, there is a ball $\mathcal{B}$ around $\mathbf{x}^*$ contained in Ω. Each $\mathbf{x} \in \mathcal{B}$ can be written as $\mathbf{x}^* + \alpha\mathbf{d}$. For small enough $\alpha > 0$ we have

$$\frac{f(\mathbf{x}^* + \alpha\mathbf{d}) - f(\mathbf{x}^*)}{\alpha^2} = \frac{1}{2}\mathbf{d}^T D^2 f(\mathbf{x}^*)\mathbf{d} + \alpha(\cdots) > 0,$$

and the result follows. For the case where $\mathbf{x}^*$ is a boundary point we have to use the convexity of Ω, as was done in the proof of Theorem 1. □

There is no need to employ an inequality of Rayleigh as was done in [3], however, note that we have used Weierstrass's theorem.

Example: derivation of the second derivative test

Consider a twice differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ which we wish to extremise. Without any constraints, the feasible region is $\Omega = \mathbb{R}$, and all points are interior points.

By Corollary 3, all extremisers must have $\nabla f(x) = 0$ (which makes them a stationary point). Notice that, since there is just one variable x, the gradient $\nabla f(x)$ is simply $\frac{df}{dx} = f'(x)$, so the claim that $\nabla f(x) = 0$ means that $f'(x) = 0$.

If $\mathbf{x}^*$ is a stationary point, then the first condition of the SOSC is satisfied. The second condition of the SOSC requires $D^2 f(x)$. As with the gradient, since there is just one variable, this is simply the second derivative $\frac{d^2 f}{dx^2} = f''(x)$. In that case, being positive definite simply means $f''(x) > 0$ (which implies a local minimiser). Similarly, being negative definite means that $f''(x) < 0$ (which implies a local maximiser).

This completely describes the second derivative test: if x_0 is a stationary point (i.e., $f'(x_0) = 0$), then:

- if $f''(x_0) > 0$, then x_0 is a local minimiser, and
- if $f''(x_0) < 0$, then x_0 is a local maximiser, and
- if $f''(x_0) = 0$, then the test is inconclusive.

Example: derivation of the test for functions in two variables

Consider a twice differentiable function $f\colon \mathbb{R}^2 \to \mathbb{R}$ which we wish to extremise. Without any constraints, the feasible region is $\Omega = \mathbb{R}^2$, and all points are interior points.

By Corollary 3, all extremisers must have $\nabla f(\mathbf{x}) = 0$ (which makes them a stationary point). Note that

$$\nabla^T f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix},$$

so this requires both partial derivatives to equal zero.

If $\mathbf{x}^*$ is a stationary point, then the first condition of the SOSC is satisfied. For the function f of two variables, we have

$$D^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix},$$

and if they are all continuous, we have $f_{xy} = f_{yx}$.

We now apply Sylvester's criterion. The principal leading minors of this matrix are f_{xx}, and $D^2 f$ itself. The determinant of $D^2 f$ is

$$f_{xx} f_{yy} - f_{xy} f_{yx} = f_{xx} f_{yy} - (f_{xy})^2.$$

Then, applying Sylvester's criterion,

- if $f_{xx} > 0$ and $f_{xx} f_{yy} - (f_{xy})^2 > 0$ then $D^2 f$ is positive definite,
- if $f_{xx} < 0$ and $f_{xx} f_{yy} - (f_{xy})^2 > 0$ then $D^2 f$ is negative definite,
- if $f_{xx} < 0$ and $f_{xx} f_{yy} - (f_{xy})^2 < 0$ then $D^2 f$ is neither positive definite nor negative definite,
- if any of f_{xx} or $f_{xx} f_{yy} - (f_{xy})^2$ equal 0, we cannot make a conclusion.

In the first case, the SOSC tells us we have a local minimiser. In the second case, the SOSC tells us we have a local maximiser. In the third case, the SONC tells us that the point is neither a local minimiser nor maximiser. Hence, we obtain the following table

$f_{xx} f_{yy} - (f_{xy})^2$ negative $\longrightarrow$	saddle point
$f_{xx} f_{yy} - (f_{xy})^2$ positive $\longrightarrow$	f_{xx} positive $\longrightarrow$ minimum
	f_{xx} negative $\longrightarrow$ maximum

To conclude, we have presented a theoretical basis for the solution of non-linear unconstrained and set-constrained problems. We have improved upon the theory presented in [3], by providing a first order sufficient condition using (local) convexity of the constrained set. Also we indicated that the SOSC can be extended to boundary points.

We reiterate here that, when the objective function is nonlinear and depends on many variables, there is no way we will be able to solve the conditions presented here. Therefore we will have to turn to iterative methods to find approximate solutions. This we do in Chapter 3, starting with search methods for functions of one variable.