TPP Calculus Review
Fall 2021
Lecture Notes*

Manon Revel

August 19, 2021

Contents

*These notes are inspired from a previous review conducted by Jesse Jenkins

1 Basic definitions

1.1 Functions

A function defines a relation between two sets, and associates each element of the initial set to exactly one element in the image set. We will investigate hereafter real functions, that take values from $\mathbb{R}$ to $\mathbb{R}$.

> Let
> $$f : \mathbb{R} \to \mathbb{R}, \\ x \mapsto y = f(x)$$
> be a function that maps each value $x \in \mathbb{R}$ to its image $y \in \mathbb{R}$.

Example:

- $f_1(x) = x$
- $f_2(x) = 3x - 2$
- $f_3(x) = \cos(x)$
- $f_4(x) = -x^2 - 8x - 12$
- $f_5(x) = x^2 + 25$
- $f_6(x) = x^3$

A function f is injective if each element in the image set has at most one pre-image in the initial set.

A function f is surjective if each element in the image set has at least one pre-image in the initial set.

A function is bijective if it is injective and surjective.

Exercise: Find which diagram represents a bijective, surjective, injective function.

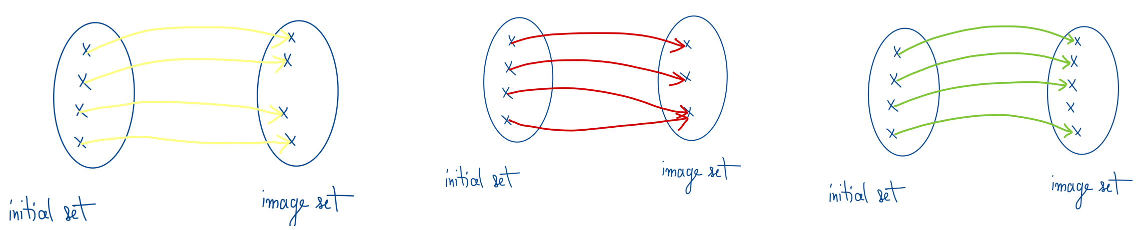


Figure 1: Bijection, Surjection and Injection

Exercise: Which of the functions above-mentioned are bijections from $\mathbb{R}$ to $\mathbb{R}$?

1.2 Linear Functions

There exists different particular functions. The simplest ones are the linear functions.

> A linear function is a polynomial function of degree 0 or 1, that is a function of the form $f(x) = ax + b$, where a is called the slope and b the intercept.

The **intercept** is the value of the function at $x = 0$. Indeed, $f(0) = a \times 0 + b = b$.

The **slope** is the incremental change in the function when x increases by one. Indeed, $f(x+1) - f(x) = a \times (x+1) + b - a \times x - b = a$.

Exercise: Which of the above-defined functions are linear? What are their slope and intercept? What does it mean for a linear function to have positive/negative slope?

1.3 Polynomial Functions

We can then generalize to polynomial functions.

A polynomial function of degree n is such that $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_i \in \mathbb{R}$, for all i in $[0, n]$.

One often encounters quadratic functions (of degree 2) and cubic equations (of degree 3).

Exercise: Which if the above-defined functions are quadratic? Cubic?

1.4 Inverse Functions

An inverse function reverses the initial function. That is, for a function $f(x) = y$ that maps each value x to a value y, the inverse function $f^{-1}(y) = x$ maps each value y to a value x. Only bijections are invertible.

Example:

- $f_1^{-1}(y) = y$
- $f_2^{-1}(y) = \frac{y+2}{3}$
- f_3 is not invertible as a function from $\mathbb{R}$ to $\mathbb{R}$; $f_3^{-1}(y) = \arccos(y)$, restricting $y \in [-1, 1]$
- f_4 is not invertible
- f_5 is not invertible as a function from $\mathbb{R}$ to $\mathbb{R}$; $f_5^{-1}(y) = \sqrt{y - 25}$, restricting $y \in [25, \infty]$
- $f_6^{-1}(y) = \sqrt[3]{y}$

2 Comparison of Linear Functions

2.1 Comparison

Let f and g be two linear functions, such that $f(x) = ax + b$ and $g(x) = cx + d$. We want to find the intervals I_- and I_+ such that $f(x) \leq g(x)$, for all $x \in I_-$ and $f(x) \geq g(x)$, for all $x \in I_+$.

One can identify these intervals analytically, $I_- = [-\infty, \frac{d-b}{a-c}]$ and $I_+ = [\frac{d-b}{a-c}, \infty]$.

At $\frac{d-b}{a-c}$, both functions are equal, provided that $a \neq c$. Otherwise, the fraction is not well-defined.

Exercise: What does it means for $a = c$? How do the functions' plots look like?

2.2 Solving Linear Systems

A linear system is a list of linear equations defined on the same set of variables, from which one shall identify the values of the variables.

2.2.1 2×2 system

For instance, a 2×2–system looks like:

$$ax - y = -b$$
$$cx - y = -d$$

which can be re-arranged to see that one actually is looking for the point at which two linear functions are equal. Yet, usually, these systems are solved using the Gauss Eliminating method. Let us look at the following example:

$$\begin{aligned} 2x + y &= 5 \\ x - 3y &= -1 \end{aligned}$$

One could re-arrange the equations to identify the linear functions, and conclude that $x = 2$ noticing that $a = -2, b = 5, c = 1/3, d = 1/3$.

Otherwise, one can eliminate one-variable by parametrizing one variable in each equation with the same number. For instance, if one multiplies the first line by 3 :

$$\begin{aligned} 3 \times (2x + y) &= 3 \times (5) \\ x - 3y &= -1, \end{aligned}$$

$3y$ emerges in both lines.

$$\begin{aligned} 6x + 3y &= 15 \\ x - 3y &= -1 \end{aligned}$$

It hence suffices to sum both lines to cancel the dependence in y.

$$6x + 3y + x - 3y = 15 - 1$$

Finally, $7x = 14$, so $x = 2$. One can find y using either of the initial equations: $y = 5 - 2 \times 2 = 1$, or $y = \frac{1+2}{3} = 1$.

Relating to what we saw in the linear algebra review, note that we can also define this system with matrices. Indeed, note that we can re-write the system as follows:

$$A * \vec{v} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix} * \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

Note that if the matrix A is invertible, then we can simply use A's inverse to conclude that x is the first element of the vector $(5, -1)^T * A^{-1}$, and y is its second element. The system hence has a solution if and only if the matrix I is invertible, which is true if and only if A is full-rank. This means for both equations are linearly independent, or, in the case of a 2×2–system, that one equation cannot be derived from the other one, or that both lines don't have the same slope!

2.2.2 $n \times n$ system

One can generalize the latter notions to an $n \times n$ system, defined as follow:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

We can re-write this system in matrix form:

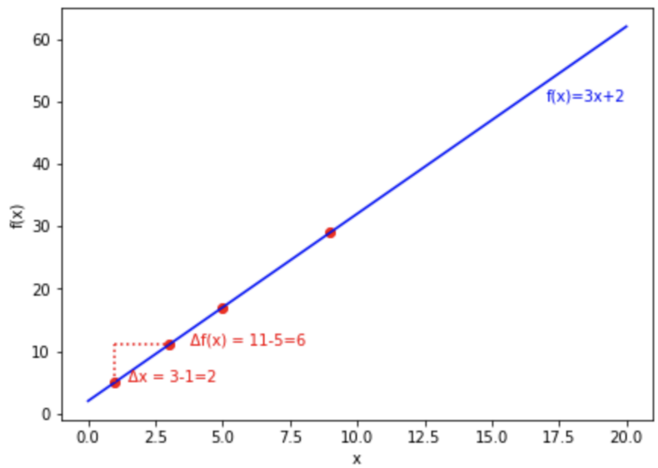
$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

3 Derivation

3.1 Slope of Linear Functions

Recall that we defined the *slope* a of a linear function $f(x) = ax + b$ as the increase when x increases by 1. We can define, more generally the slope as $a = \frac{f(x_2)-f(x_1)}{x_2-x_1} = \frac{\Delta f(x)}{\Delta x}$.

Hence, we can compute the slope of a linear function through any pair of coordinates.



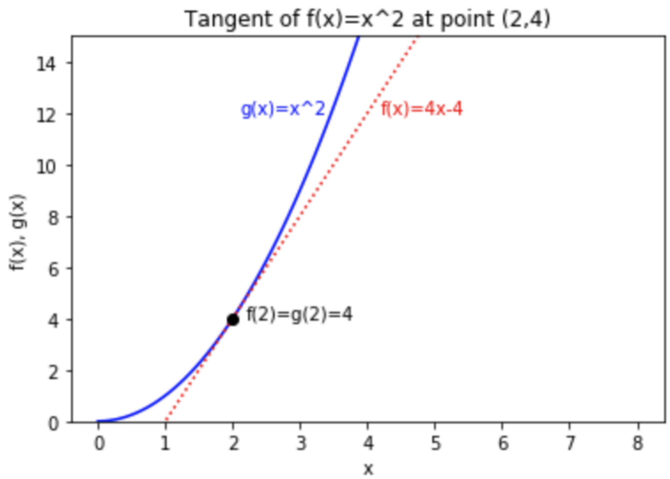
Exercise: Let f be a linear function, and let $\{(1,5),(3,11),(5,17),(9,29)\}$ be points such that $f(x) = y$. What are the slope and the intercept of f?

The slope a is also called the **derivative** of the function f at any point $x \in \mathbb{R}$. For a linear function, the derivative is the same at any point. Yet, we can generalize the notion of derivative to non-linear functions, as the slope of the function at a specific point.

3.2 Slope of Non-linear Functions

The slope of a non-linear function at x is the relative change of a function around x. The function can be described as locally linear at a point x, and the slope of the function at x is hence the slope of the linear function that approximate the function at that point. The latter function is called the **tangent** of the function at the point x.

Example: Let $g(x) = x^2$ be a quadratic function, one can find that the tangent of g at $(2,4)$ is the linear function $f(x) = 4x - 4$, shown in the following figure.



Now, note that the slope of the tangent of g at x increases with x, for $x > 0$. This indicates that the function g increases on $x > 0$.

3.3 Derivatives

Formally, let f be a function, then the slope of f at a given point x_0 is defined as:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

It is hence f's infinitesimal change around x_0. We write indifferently $f'(x) = \frac{df(x)}{dx}$.

$f'(x)$ is a also a function, of which we can take the derivative! We can the second derivative of f the derivative of f', and note f''. Formally, $f''(x) = \frac{d}{dx} f'(x) = \frac{d^2}{dx^2} f(x)$.

Example: Let $g(x) = x^2$ and $x_0 \in \mathbb{R}$,

$$\begin{aligned} g'(x_0) &= \lim_{h \to 0} \frac{g(x_0 + h) - g(x_0)}{h} \\ &= \lim_{h \to 0} \frac{(x_0 + h)^2 - x_0^2}{h} \\ &= \lim_{h \to 0} \frac{2x_0 h + h^2}{h} \\ &= \lim_{h \to 0} 2x_0 + h \\ &= 2x_0 \end{aligned}$$

so the derivative of $g(x) = x^2$ is $g'(x) = 2x$.

Now, let (x_1, x_1^2) be a point that belongs to g, let us find the equation of the equation of g's tangent f at x_1. We know that the slope of the tangent is $2x_1$ so the equation is of the form $f(x) = 2x_1 x + b$. Further, the line crosses the point (x_1, x_1^2), so $b = x_1^2 - 2x_1^2 = -x_1^2$. Finally, the tangent f of g at the given point x_1 is $f(x) = 2x_1 x - x_1^2$.

3.4 Monotonicity

Functions can be monotonic, in that they only increase or decrease in certain intervals. $f(x) = x$ is strictly increasing on $\mathbb{R}$, while $f(x) = -x$ is strictly decreasing on $\mathbb{R}$. Alternatively, sinusoidal functions increase on certain intervals and decrease on other intervals.

A function f increases on the interval $\mathcal{I}$ if $\forall x_1, x_2 \in \mathcal{I}$ such that $x_1 < x_2, f(x_1) \leq f(x_2)$.
A function f decreases on the interval $\mathcal{I}$ if $\forall x_1, x_2 \in \mathcal{I}$ such that $x_1 < x_2, f(x_1) \geq f(x_2)$.

Exercise: Name the intervals on which $f(x) = \cos(x)$ increases and decreases.

Note that if the derivative $f'(x_0) > 0$, it means that locally $f(x_0 + h) - f(x_0) > 0$, so the function is locally increasing.

More generally:

If $\forall x \in \mathcal{I}, f'(x) \geq 0$, f **is increasing** on $\mathcal{I}$.
Vice versa, $\forall x \in \mathcal{I}, f'(x) \leq 0$, f **is decreasing** on $\mathcal{I}$. If the inequalities are strict, we say that the function is strictly monotone.

Exercise:

- Check the intervals you obtained for $f(x) = \cos(x)$'s monotonicity using its derivative and the fact that $\cos(x) = \sin(\pi/2 - x)$.
- Can a *continuous* bijective function increase on a interval and decrease on another one?

3.5 Derivation in Practice

Let us state some derivatives' formulas for derivatives. Let $c \in \mathbb{R}$ be a constant.

$\frac{d}{dx}(c) = 0$

$\frac{d}{dx}(x) = 1$

$\forall n \in \mathbb{N}, \frac{d}{dx}(x^n) = nx^{n-1}$

$\frac{d}{dx}(ax + b) = a$

$\frac{d}{dx}(e^x) = e^x$

$\frac{d}{dx}(a^x) = a^x \ln(a)$

$\frac{d}{dx}(\ln x) = \frac{1}{x}$

$\frac{d}{dx}(\log_a x) = \frac{1}{x \log(a)}$

$\frac{d}{dx}(\sqrt{x}) = \frac{1}{2\sqrt{x}}$

$\frac{d}{dx}(\frac{1}{x}) = -\frac{1}{x^2}$

$\frac{d}{dx}(\sin(x)) = \cos(x)$

$\frac{d}{dx}(\cos(x)) = -\sin(x)$

$\frac{d}{dx}(\tan(x)) = \frac{1}{\cos^2(x)}$

$\frac{d}{dx}(\sin^{-1}(x)) = \frac{1}{\sqrt{1-x^2}}$

$\frac{d}{dx}(\cos^{-1}(x)) = -\frac{1}{\sqrt{1-x^2}}$

$\frac{d}{dx}(\tan^{-1}(x)) = \frac{1}{1+x^2}$

Now, let us state some formulas for the derivatives of combined functions. let $u(x), v(x)$ be two functions.

The derivative of a sum is the sum of the derivative: $\frac{d}{dx}(u(x) \pm v(x)) = \frac{du(x)}{dx} \pm \frac{dv(x)}{dx}$

The derivative of a function multiplied by a scalar is: $\frac{d}{dx}(cu(x)) = c\frac{du(x)}{dx}$

The derivative of a product is: $\frac{d}{dx}(u(x) \times v(x)) = u(x)v'(x) + u'(x)v(x)$

The derivative of a quotient is: $\frac{d}{dx}(\frac{u(x)}{v(x)}) = \frac{v(x)u'(x) - u(x)v'(x)}{v^2(x)}$

The derivative of a composition is: $\frac{d}{dx}(u(v(x))) = v'(x) \times u'(v(x))$

Exercise 1: Are the first ten above-mentioned functions increase or decrease? When needed, state the conditions required to answer.

Exercise 2: Find the derivatives of the following quantities:

- $f_1(x) = 2x^{12} + 4x^4 - x^2 + x/3 - 1 + \sin(x)$
- $f_2(x) = 3e^x$
- $f_3(x) = \frac{1}{x}$
- $f_4(x) = \frac{1}{x^2}$
- $f_5(x) = \cos(x)\tan(x)$
- $f_6(x) = \sin(\frac{1}{\ln(x)})$

4 Local and Global Optima

4.1 When the Derivative Equals 0

We have seen that the derivative gives us an idea of how the function varies locally. Now, what does it mean for a derivative to be equal to 0?

In the figure below, one sees three functions in the first raw, as well as their first and second derivatives in the second and third raws respectively.

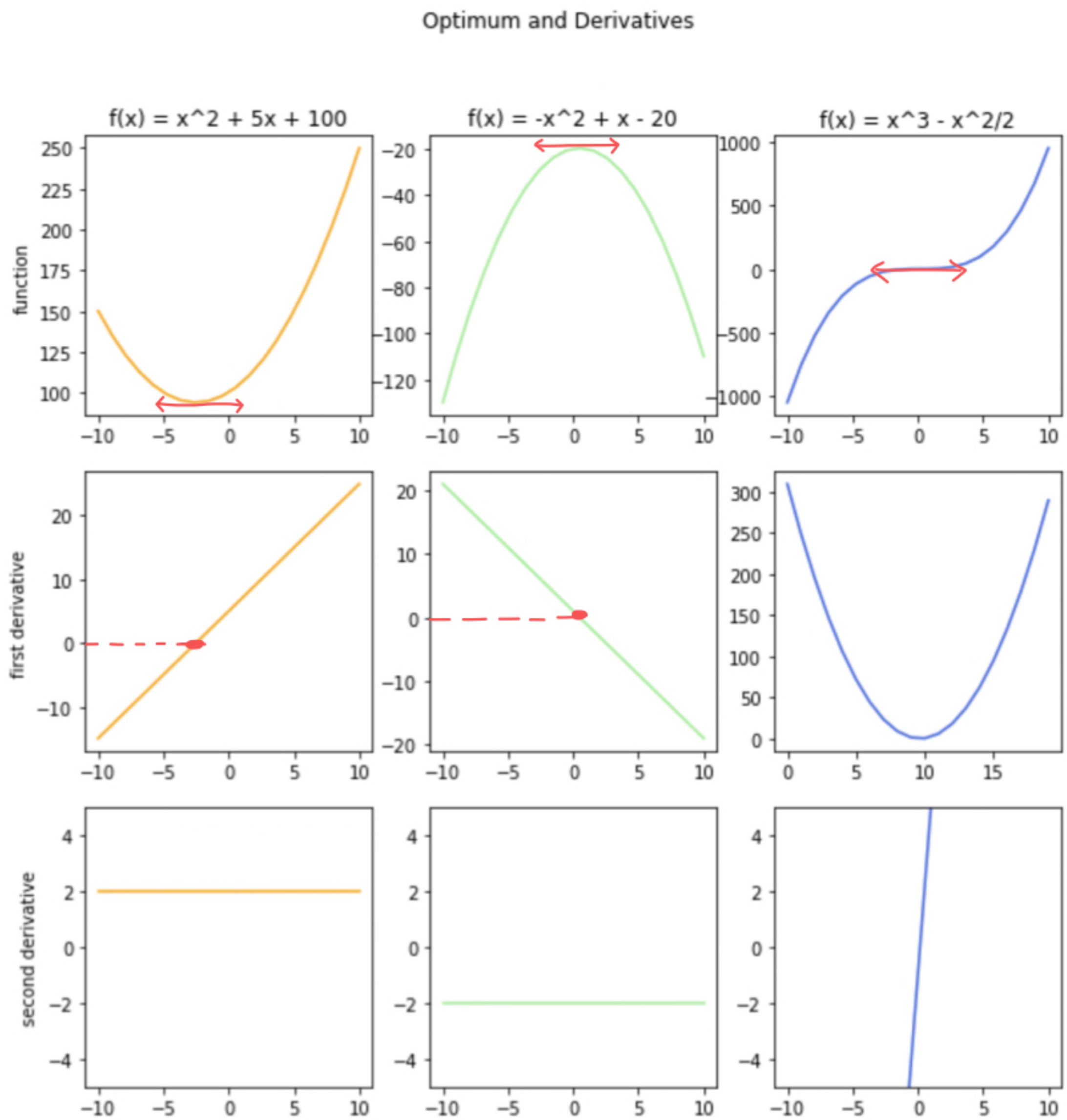
We see that for the first two functions, $f(x) = x^2 + 5x + 100$ and $f(x) = -x^2 + x - 20$, the derivatives reach 0 at a point that is the function's optimum.

Let us focus on $f(x) = x^2 + 5x + 100$. Before the point at which the derivative is null, the derivative is negative – the function decreases – and after that point, the derivative is positive – the function increases. The point at which the derivative is null is hence the *critical point* at which the function is *minimal.* Note that the second derivative of f at the critical point is 2, a positive number.

Let us focus on $f(x) = -x^2 + x - 20$. Conversely, before the point at which the derivative is null, the derivative is positive – the function increases – and after that point, the derivative is negative – the function decreases. The point at which the derivative is null is hence the *critical point* at which the function is *maximal.* Note that the second derivative of f at the critical point is -2, a negative number.

Finally, we see that for $f(x) = x^3 - \frac{x^2}{2}$, the function keeps increasing but the rate at which it does slow down before and after the *critical point* at which the local slope is null. This point is called an *inflection point.* Note that the second derivative of f at the critical point is 0.

These are the three cases in which the derivative of a function is null.



4.2 Optimum and Convexity

Critical values of interests in functional analysis are typically a function's optimal values:

> Let $f(x)$ be a function. We say that $f(x)$ is reaches its **maximum** $f(x^*)$ at x^* if $\forall x \in \mathbb{R}$, $f(x^*) \geq f(x)$. Conversely, we say that $f(x)$ is reaches its **minimum** $f(x^*)$ at x^* if $\forall x \in \mathbb{R}$, $f(x^*) \leq f(x)$.

In fact, the first and second derivatives are all you need to characterize the type of critical point under study, and importantly to characterize the optima of a function. Namely:

Let $f(x)$ be a function whose first and second derivative are $f'(x)$ and $f''(x)$ respectively, and let $(x_0, f(x_0))$ be a point such that $f'(x_0) = 0$, the derivative is null at that point. Now,

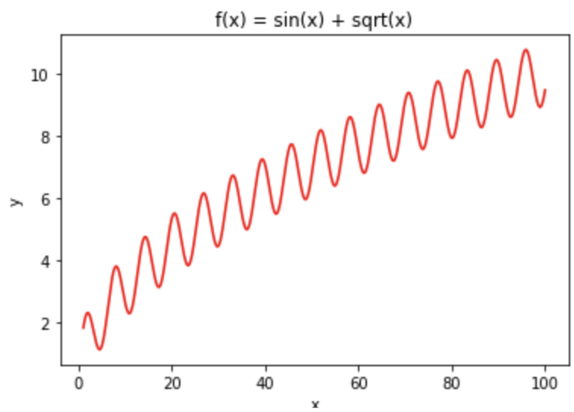
- If $f''(x_0) > 0$, the function is **minimal** at x_0.
- If $f''(x_0) < 0$ the function is **maximal** at x_0.

Further, we say that a function is **convex** on an interval $\mathcal{I}$ if $\forall x \in \mathcal{I}, f''(x) \geq 0$. Conversely, we say that a function is **concave** on an interval $\mathcal{I}$ if $\forall x \in \mathcal{I}, f''(x) \leq 0$.

4.3 Global or Local Optimum?

Note finally that one function may have multiple optima. We say that x^* is a global optimum if $f(x^*)$ is more extreme than $f(x)$ for any real x. On the other hand, we say that x^* is a local optimum if $f(x^*)$ is more extreme than $f(x)$ for any x in a given interval.

Example: Sinusoidal functions are typical examples of functions with multiple optima. We can focus on certain intervals and identify local optima on those. Let $f(x) = \sin(x) + \sqrt{x}$, as shown below, this function has multiple local maxima (points at which the first derivative is null) but non of these are global (in fact, the function increases to infinity).



Exercise: Find whether the function has optimal points with the least possible calculations. Say whether it is maximum or minimum, and whether it is a local or global optimum.

- $f_1(x) = (x - 4)(2x - 1)$
- $f_2(x) = -(x + 4)(2x - 1)$
- $f_3(x) = (x - 2)(x - 1)(x + 6)$
- $f_4(x) = x^5 - 3x + 5$
- $f_5(x) = \frac{100}{x} + 25x$

5 Optimization

One might aim at maximize a function, while being constrained by a set of conditions.

A shoemaker sells two types of pair of shoes: the first type $T1$ is sold at \$50 and is made out of one unit of leather; the second type $T2$ is sold at \$75 and is made out of two units of leather. Each shoe type requires a pair of shoelace. Each shoelaces pair costs \$5 and a unit of leather costs \$25. The shoemaker wants to use all 100 leather units – yet only has 70 shoelace pairs at her disposal – while maximizing her profit.

5.1 Optimization Problem

This problem can be formulated as an optimization problem, in which one aims at optimizing an objective function submitted to a set of constraints.

5.1.1 Variables

First, let's define the variables associated with the problem: these are the numbers of shoes pairs of each type one shall produced, let call them T_1 and T_2.

5.1.2 Objective Function

The shoemaker wants to maximising her total profit, that is her net gain after producing and selling the shoes. Assuming that all produced shoes would be sold, the profit is $50T_1 + 75T_2 - 25T_1 - 2 * 25T_2 - 5T_1 - 5T_2 = 20(T_1 + T_2)$. This is the objective function that we aim to maximize. As such, one would take both variables to be infinite! Here is where the constraints intervene.

5.1.3 Constraints

The shoemaker wants to use all her leather units, so that $T_1 + 2T_2 = 100$. Further, there is a total of 70 shoelace, so $T_1 + T_2 \leq 70$. These two limitations are the problem's constraints.

5.1.4 Variable Type

Finally, note that the variables are bounded to be positive, and integers.

We can hence formulate the optimization problem as follow:

$$\begin{aligned} \underset{x,y}{\text{maximize}} \quad & x + y \\ \text{subject to} \quad & x + 2y = 100 \\ & x + y \leq 70 \\ & (x, y) \in \mathbb{N} \end{aligned}$$

More generally, we can define an optimization problem defining the variables x_i for $i \in [1, n]$, and objective function $f(x_1, ..., x_n)$ and constraints.

$$\begin{aligned} \underset{\vec{x}}{\text{minimize}} \quad & f(\vec{x}) \\ \text{subject to} \quad & h_i(\vec{x}) = 0, \; i = 1, \ldots, m \\ & g_i(\vec{x}) \leq 0, \; i = 1, \ldots, r. \end{aligned}$$

5.2 Lagrange Multiplier

Let us consider a problem only with m equality constrains. To solve such an optimization problem, we define the **Lagrangian function**

$$\begin{aligned} &\mathcal{L}: \mathbb{R}^{n+m} \to \mathbb{R}, \\ &(\vec{x}, \vec{\lambda}) \mapsto \mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \textstyle\sum_{i=1}^{m} \lambda_i h_i(\vec{x}) \end{aligned}$$

where $\vec{\lambda} = (\lambda_1, \cdots, \lambda_m)^T$ is the vector of the Lagrangian multipliers corresponding to the equality constraints.

We set the partial derivatives of $\mathcal{L}$ with respect to each of the $n + m$ variables to 0, creating an $n + m \times n + m$ system of equations (potentially linear)! Solving it provides us with the optimal values of the variables in $\vec{x} = (x_1, \cdots, x_n)^T$.

Exercise: Solve the following optimization problem:

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & f(x,y) = xe^{-y} \\ \text{subject to} \quad & 2x - y = 10 \end{aligned}$$