Energy Systems, Summer Semester 2026
Lecture 7: Optimisation

Prof. Tom Brown, Philipp Glaum
Department of Digital Transformation in Energy Systems, Institute of Energy Technology, TU Berlin

Table of Contents

Technische Universität Berlin

Optimisation: Motivation

What to do about variable renewables?

Backup energy costs money and may also cause CO_2 emissions.

Curtailing renewable energy is also a waste.

We consider **four options** to deal with variable renewables:

1. Smoothing stochastic variations of renewable feed-in over **larger areas**, e.g. the whole of European continent.
2. Using **storage** to shift energy from times of surplus to deficit.
3. **Shifting demand** to different times, when renewables are abundant.
4. Consuming the electricity in **other sectors**, e.g. transport or heating.

Optimisation in energy networks is a tool to assess these options.

Why optimisation?

In the energy system we have lots of **degrees of freedom**:

1. Power plant and storage dispatch
2. Renewables curtailment
3. Dispatch of network elements (e.g. High Voltage Direct Current (HVDC) lines)
4. Capacities of everything when considering investment

but we also have to respect **physical constraints**:

1. Meet energy demand
2. Do not overload generators or storage
3. Do not overload network

and we want to do this while **minimising costs**. Solution: **optimisation**.

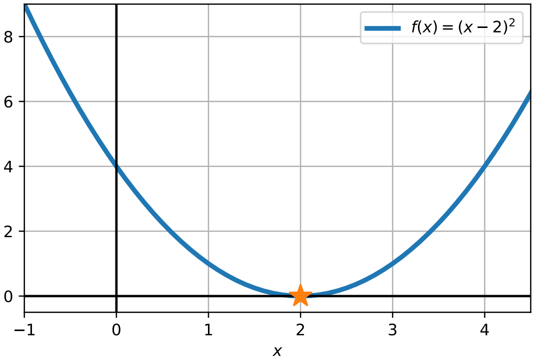
Optimisation: Introduction

Simplest 1-d optimisation problem

Consider the following problem. We have a function $f(x)$ of one variable $x \in \mathbb{R}$

$$f(x) = (x-2)^2$$

Where does it reach a minimum? School technique: find stationary point $\frac{df}{dx} = 2(x-2) = 0$, i.e. minimum at $x^* = 2$ where $f(x^*) = 0$.

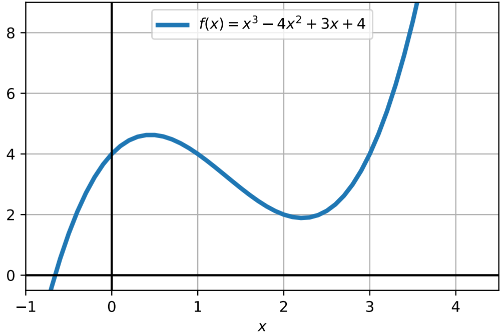


Simplest 1-d optimisation problem

Consider the following problem. We have a function $f(x)$ of one variable $x \in \mathbb{R}$

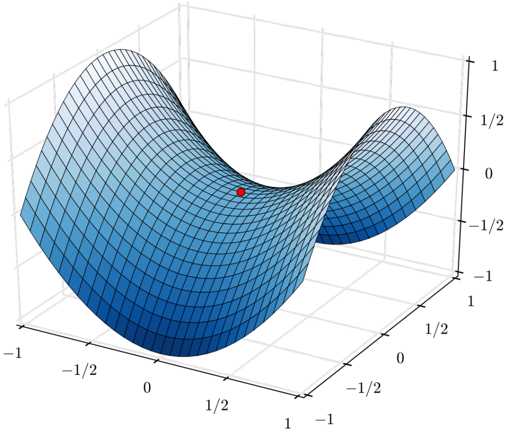
$$f(x) = x^3 - 4x^2 + 3x + 4$$

Where does it reach a minimum? School technique fails since has two stationary points, one local minimum and local maximum; must check 2nd derivative for minimum/maximum. Also: function is not bounded as $x \to -\infty$. No solution!



Beware saddle points in higher dimensions

Some functions have **saddle points** with zero derivative in all directions (stationary points) but that are neither maxima nor minima, e.g. $f(x, y) = x^2 - y^2$ at $(x, y) = (0, 0)$.



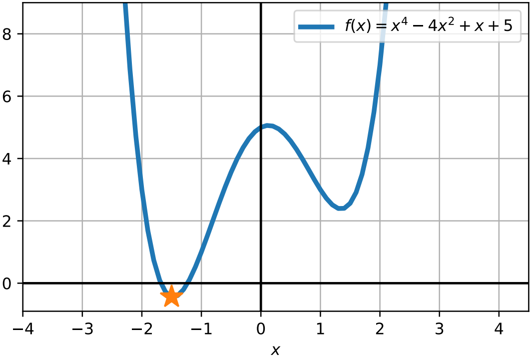
Source: Wikipedia

Simplest 1-d optimisation problem

Consider the following problem. We have a function $f(x)$ of one variable $x \in \mathbb{R}$

$$f(x) = x^4 - 4x^2 + x + 5$$

Where does it reach a minimum? Now two separate local minima. Function is **not convex** downward. This is a problem for algorithms that only search for minima locally.



Simplest 1-d optimisation problem with constraint

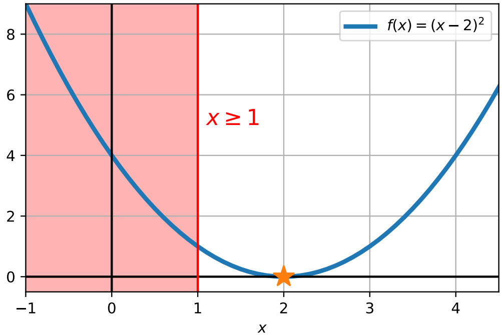
Consider the following problem. We minimise a function of one variable $x \in \mathbb{R}$

$$\min_x (x-2)^2$$

subject to a constraint

$$x \geq 1$$

The constraint has **no effect** on the solution. It is **non-binding**.



Simplest 1-d optimisation problem with constraint

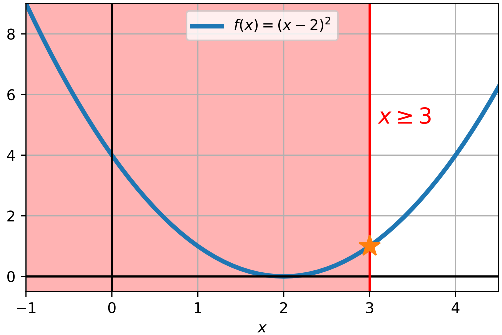
Consider the following problem. We minimise a function of one variable $x \in \mathbb{R}$

$$\min_x (x-2)^2$$

subject to a constraint

$$x \geq 3$$

Now the constraint is **binding** and is **saturated** at the optimum $x^* = 3$.



Simple 2-d optimisation problem

Consider the following problem. We have a function $f(x, y)$ of two variables $x, y \in \mathbb{R}$

$$f(x, y) = 3x$$

and we want to find the maximum of this function in the $x - y$ plane

$$\max_{x,y \in \mathbb{R}} f(x, y)$$

subject to the following constraints

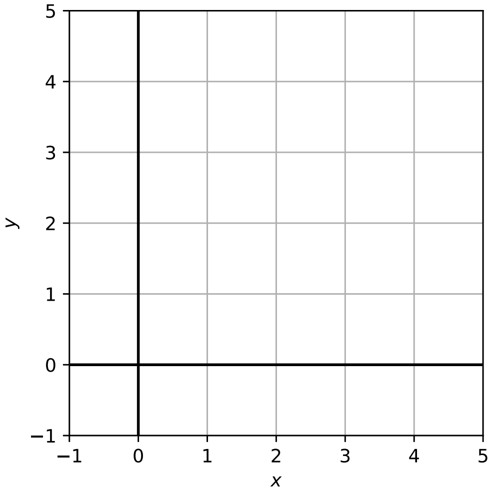
$$x + y \leq 4 \tag{1}$$

$$x \geq 0 \tag{2}$$

$$y \geq 1 \tag{3}$$

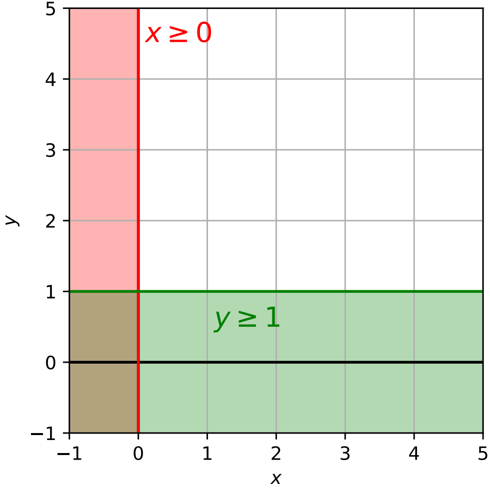
Simple 2-d optimisation problem

Consider $x - y$ plane of our variables:

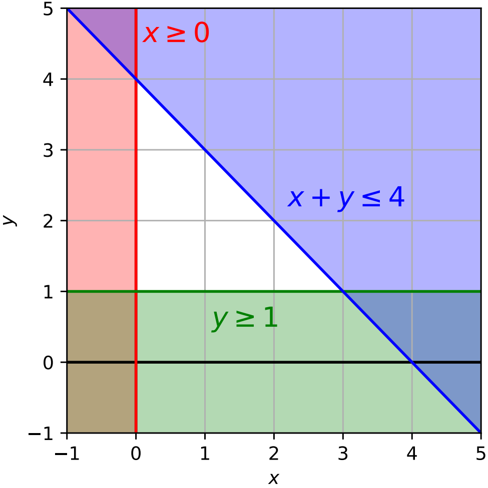


Simple 2-d optimisation problem

Add constraints (2) and (3):

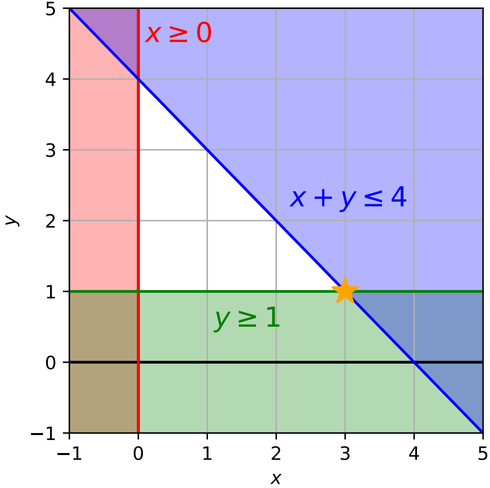


Add constraint (1). In this allowed space (white area) what is the maximum of $f(x, y) = 3x$?



Simple 2-d optimisation problem

$f(x, y) = 3x$ maximised at $x^* = 3, y^* = 1, f(x^*, y^*) = 9$:



Simple 2-d optimisation problem

Consider the following problem. We have a function $f(x, y)$ of two variables $x, y \in \mathbb{R}$

$$f(x, y) = 3x$$

and we want to find the maximum of this function in the $x - y$ plane

$$\max_{x,y \in \mathbb{R}} f(x, y)$$

subject to the following constraints

$$x + y \leq 4 \tag{4}$$

$$x \geq 0 \tag{5}$$

$$y \geq 1 \tag{6}$$

Optimal solution: $x^* = 3, y^* = 1, f(x^*, y^*) = 9.$

NB: We would have gotten the same solution if we had removed the 2nd constraint - it is **non-binding**.

Another simple optimisation problem

We can also have equality constraints. Consider the maximum of this function in the $x - y - z$ space

$$\max_{x,y,z \in \mathbb{R}} f(x, y, z) = (3x + 5z)$$

subject to the following constraints

$$\begin{aligned} x + y &\leq 4 \\ x &\geq 0 \\ y &\geq 1 \\ z &= 2 \end{aligned}$$

Another simple optimisation problem

We can also have equality constraints. Consider the maximum of this function in the $x - y - z$ space

$$\max_{x,y,z \in \mathbb{R}} f(x, y, z) = (3x + 5z)$$

subject to the following constraints

$$\begin{aligned} x + y &\leq 4 \\ x &\geq 0 \\ y &\geq 1 \\ z &= 2 \end{aligned}$$

Optimal solution: $x^* = 3, y^* = 1, z^* = 2, f(x^*, y^*, z^*) = 19.$

[This problem is **separable**: can solve for (x, y) and (z) separately.]

Energy system mapping to an optimisation problem

This optimisation problem has the same basic form as our energy system considerations:

Objective function to minimise	↔	**Minimise total costs**
Optimisation variables	↔	**Physical degrees of freedom (power plant dispatch, etc.)**
Constraints	↔	**Physical constraints (overloading, etc.)**

Have to meet annual electricity demand of 500 TWh_{el}/a. Have two generators:

Generator	Capacity limit GW_{el}	Yearly limit TWh_{el}/a	efficiency	fuel cost €$/MWh_{th}$	specific emissions tCO_2/MWh_{th}
gas	40	350	0.6	20	0.2
coal	34	300	0.4	8	0.3

- What is the minimum cost solution (considering only fuel costs)?
- What is the minimum cost solution if we restrict emissions to 300 $MtCO_2/a$?
- What about 250 $MtCO_2/a$?

2 variables: x_1 for yearly electricity generation from gas, x_2 for generation from coal.

Specific electricity costs are given by fuel cost divided by efficency. For gas

$$o_1 = \frac{20\,€/\mathrm{MWh_{th}}}{0.6} = 33\,€/\mathrm{MWh_{el}}$$

For coal

$$o_2 = \frac{8\,€/\mathrm{MWh_{th}}}{0.4} = 20\,€/\mathrm{MWh_{el}}$$

So the cheapest option is to produce as much from coal as possible. But there are capacity constraints!

$\Rightarrow$ Generate maximum from coal (300 $\mathrm{TWh_{el}}$/a), rest from gas (200 $\mathrm{TWh_{el}}$/a).

Consume 750 $\mathrm{TWh_{th}}$/a of coal and 333 $\mathrm{TWh_{th}}$/a of gas $\Rightarrow$ emissions of $225 + 67 = 292$ $MtCO_2$/a.

Simple energy system example

The objective function minimises total electricity generation costs:

$$\min_{x_1, x_2} f(x_1, x_2) = o_1 x_1 + o_2 x_2$$

such that we meet the yearly demand:

$$x_1 + x_2 = 500$$

and we respect the capacity constraints:

$$0 \leq x_1 \leq 350$$

$$0 \leq x_2 \leq 300$$

as well as the emissions limit:

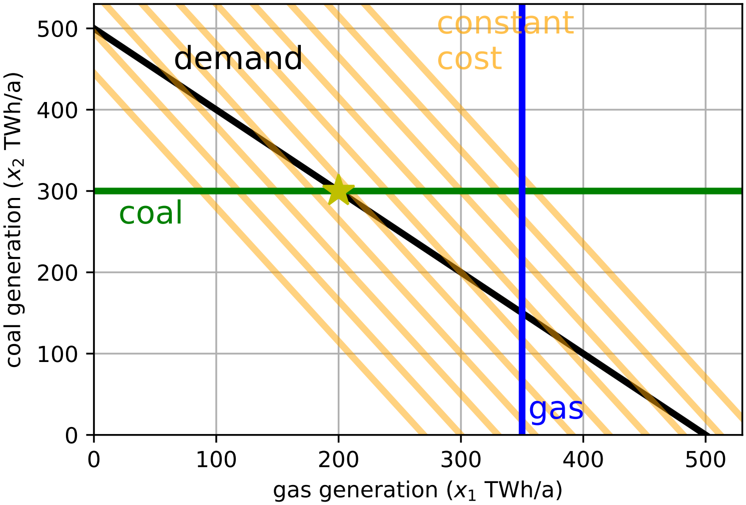
$$\frac{e_1}{\eta_1} x_1 + \frac{e_2}{\eta_2} x_2 \leq 300$$

where e_i is the specific emissions and η_i is the efficiency.

Optimal solution: $x_1^* = 200, x_2^* = 300$.

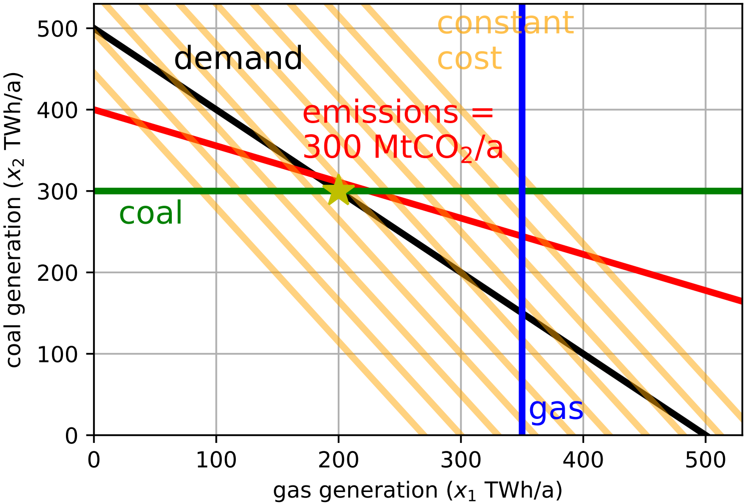
Simple energy system example

The optimal solution is given by $x_1^* = 200, x_2^* = 300$.



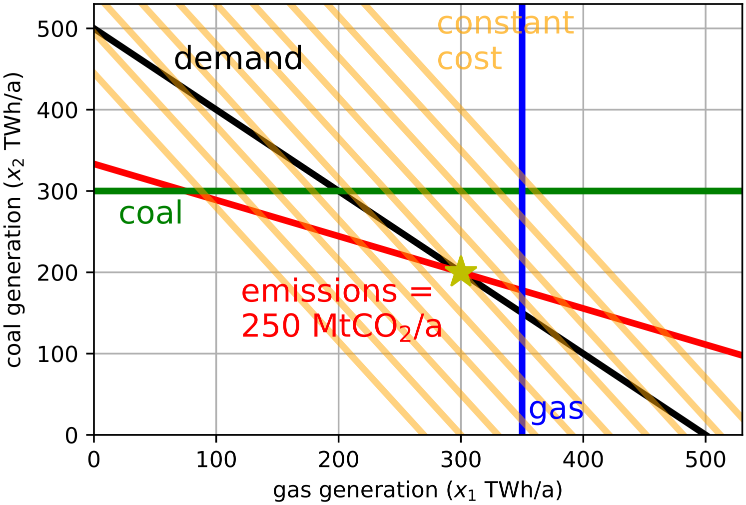
Simple energy system example

An emissions cap of ≤ 300 $MtCO_2/a$ is non-binding - the solution stays the same.



Simple energy system example

If we restrict emissions from 300 to 250 $MtCO_2/a$, the solution shifts to $x_1^* = 300, x_2^* = 200$.



Optimisation: Theory

Optimisation problem

We have an **objective function** $f : \mathbb{R}^k \to \mathbb{R}$

$$\max_x f(x)$$

$[x = (x_1, \ldots x_k)]$ subject to some **constraints** within $\mathbb{R}^k$:

$$\begin{array}{cccc} g_i(x) = c_i & \leftrightarrow & \lambda_i & i = 1, \ldots n \\ h_j(x) \leq d_j & \leftrightarrow & \mu_j & j = 1, \ldots m \end{array}$$

λ_i and μ_j are the **Karush-Kuhn-Tucker (KKT) multipliers** (basically Lagrange multipliers) we introduce for each constraint equation. Each one measures the change in the objective value of the optimal solution obtained by relaxing the constraint by a small amount. Informally $\lambda_i \sim \frac{\partial f}{\partial c_i}$ and $\mu_j \sim \frac{\partial f}{\partial d_j}$ at the optimum x^*. They are also known as the **shadow prices** of the constraints.

Feasibility

The space $X \subset \mathbb{R}^k$ which satisfies

$$
\begin{array}{cccc}
g_i(x) = c_i & \leftrightarrow & \lambda_i & i = 1, \ldots n \\
h_j(x) \leq d_j & \leftrightarrow & \mu_j & j = 1, \ldots m
\end{array}
$$

is called the **feasible space**.

It will have dimension lower than k if there are independent equality constraints.

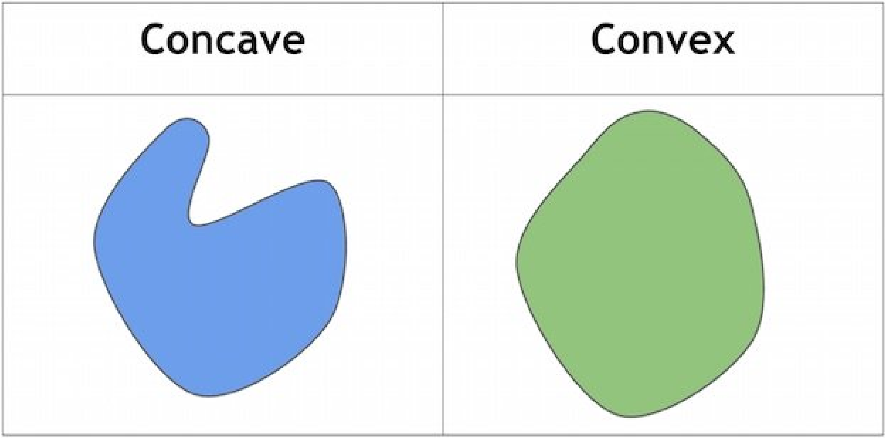
It may also be empty (e.g. for $k = 1$, $x \geq 1, x \leq 0$ in $\mathbb{R}^1$), in which case the optimisation problem is called **infeasible**.

It can be **convex** or **non-convex**.

If all the constraints are affine, then the feasible space is a convex polytope (multi-dimensional polygon).

Convexity means fast polynomial algorithms

If the feasible space is **convex** it is much easier to search, since for a convex objective function we can keep looking in the direction of improving objective function without worrying about getting stuck in a local maximum.



Lagrangian

We now study the **Lagrangian function**

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_i \lambda_i \left[g_i(x) - c_i\right] - \sum_j \mu_j \left[h_j(x) - d_j\right]$$

We've built this function using the variables λ_i and μ_j to better understand the optimal solution of $f(x)$ given the constraints.

The stationary points of $\mathcal{L}(x, \lambda, \mu)$ tell us important information about the optima of $f(x)$ given the constraints.

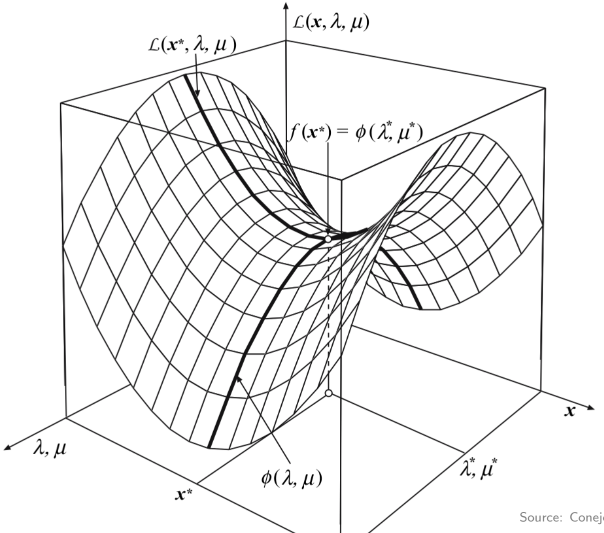
[It is entirely analogous to the physics Lagrangian $L(x, \dot{x}, \lambda)$ except we have no explicit time dependence $\dot{x}$ and we have additional constraints which are inequalities.]

We can already see that if $\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$ then the equality constraint $g_i(x) = c$ will be satisfied.

[Beware: $\pm$ signs appear differently in literature, but have been chosen here such that $\lambda_i = \frac{\partial \mathcal{L}}{\partial c_i}$ and $\mu_j = \frac{\partial \mathcal{L}}{\partial d_j}$.]

Optimum is a saddle point of the Lagrangian

The stationary point of $\mathcal{L}$ is a saddle point in (x, λ, μ) space (here minimising $f(x)$):



Source: Conejo et al, "Decomposition Techniques" (2006)

KKT conditions

The **Karush-Kuhn-Tucker (KKT) conditions** are necessary conditions that an optimal solution x^*, μ^*, λ^* always satisfies (up to some regularity conditions):

1. **Stationarity**: For $\ell = 1, \ldots k$

$$\frac{\partial \mathcal{L}}{\partial x_\ell} = \frac{\partial f}{\partial x_\ell} - \sum_i \lambda_i^* \frac{\partial g_i}{\partial x_\ell} - \sum_j \mu_j^* \frac{\partial h_j}{\partial x_\ell} = 0$$

2. **Primal feasibility**:

$$g_i(x^*) = c_i$$
$$h_j(x^*) \leq d_j$$

3. **Dual feasibility**: $\mu_j^* \geq 0$
4. **Complementary slackness**: $\mu_j^*(h_j(x^*) - d_j) = 0$

Complementarity slackness for inequality constraints

We have for each inequality constraint

$$\mu_j^* \geq 0$$
$$\mu_j^*(h_j(x^*) - d_j) = 0$$

So **either** the inequality constraint is binding

$$h_j(x^*) = d_j$$

and we have $\mu_j^* \geq 0$.

Or the inequality constraint is NOT binding

$$h_j(x^*) < d_j$$

and we therefore MUST have $\mu_j^* = 0$.

If the inequality constraint is non-binding, we can remove it from the optimisation problem, since it has no effect on the optimal solution.

Nota Bene

1. The KKT conditions are necessary conditions for an optimal solution, but are only **sufficient** for optimality of the solution under certain conditions, e.g. for problems with convex objective, convex differentiable inequality constraints and affine equalities constraints. For linear problems, KKT is sufficient.

2. The variables x_ℓ are often called the **primary variables**, while (λ_i, μ_j) are the **dual variables**.

3. Since at the optimal solution we have $g_i(x^*) = c_i$ for equality constraints and $\mu_j^*(h_j(x^*) - d_j) = 0$, we have

$$\mathcal{L}(x^*, \lambda^*, \mu^*) = f(x^*)$$

How we will use the KKT conditions

Usually we will have enough constraints to determine the k values x_ℓ^* for $\ell = 1, \dots k$ uniquely, i.e. k independent constraints will be binding and the objective function is never constant along any constraint.

We will use the KKT conditions, primarily stationarity, to determine the values of the k KKT multipliers for the independent binding constraints.

Dimensionality check: we need to find k KKT multipliers and we have k equations from stationarity to find them. Good!

The remaining KKT multipliers are either zero (for non-binding constraints) or dependent on the k independent KKT multipliers in the case of dependent constraints.

(There are also degenerate cases where the optimum is not at a single point, where things will be more complicated, e.g. when objective function is constant along a constraint.)

Return to simple optimisation problem

We want to find the maximum of this function in the $x - y$ plane

$$\max_{x,y \in \mathbb{R}} f(x,y) = 3x$$

subject to the following constraints (now with KKT multipliers)

$$\begin{aligned} x + y \leq 4 \quad & \leftrightarrow \quad \mu_1 \\ -x \leq 0 \quad & \leftrightarrow \quad \mu_2 \\ -y \leq -1 \quad & \leftrightarrow \quad \mu_3 \end{aligned}$$

We know the optimal solution in the **primal variables** $x^* = 3, y^* = 1, f(x^*, y^*) = 9$.

What about the **dual variables** μ_i?

Since the second constraint is not binding, by complementarity $\mu_2^*(-x^* - 0) = 0$ we have $\mu_2^* = 0$. To find μ_1^* and μ_3^* we have to do more work.

Simple problem with KKT conditions

We use stationarity for the optimal point:

$$0 = \frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} - \sum_i \lambda_i^* \frac{\partial g_i}{\partial x} - \sum_j \mu_j^* \frac{\partial h_j}{\partial x} = 3 - \mu_1^* + \mu_2^*$$
$$0 = \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial f}{\partial y} - \sum_i \lambda_i^* \frac{\partial g_i}{\partial y} - \sum_j \mu_j^* \frac{\partial h_j}{\partial y} = -\mu_1^* + \mu_3^*$$

From which we find:

$$\mu_1^* = 3 - \mu_2^* = 3$$
$$\mu_3^* = \mu_1^* = 3$$

Check interpretation: $\mu_j = \frac{\partial \mathcal{L}}{\partial d_j}$ with $d_j \to d_j + \varepsilon$.

Simple problem with KKT conditions: Check interpretation

Check interpretation of $\mu_1^* = 3$ by shifting constant d_1 for first constraint by ε and solving:

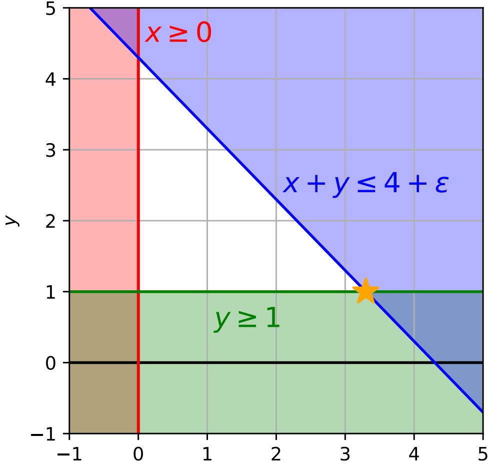
$$\max_{x,y \in \mathbb{R}} f(x,y) = 3x$$

subject to the following constraints

$$x + y \leq 4 + \varepsilon \quad \leftrightarrow \quad \mu_1$$
$$-x \leq 0 \quad \leftrightarrow \quad \mu_2$$
$$-y \leq -1 \quad \leftrightarrow \quad \mu_3$$

$f(x,y) = 3x$ maximised at $x^* = 3+\varepsilon, y^* = 1, f(x^*,y^*) = 9+3\varepsilon$.

$d_1 \to d_1 + \varepsilon$ causes optimum to shift $f(x^*,y^*) \to f(x^*,y^*) + 3\varepsilon$. Consistent with $\mu_1^* = 3$.



Return to another simple optimisation problem

We want to find the maximum of this function in the $x - y - z$ space

$$\max_{x,y,z\in\mathbb{R}} f(x,y,z) = 3x + 5z$$

subject to the following constraints (now with KKT multipliers)

$$\begin{aligned} x + y &\leq 4 &\leftrightarrow&\quad \mu_1 \\ -x &\leq 0 &\leftrightarrow&\quad \mu_2 \\ -y &\leq -1 &\leftrightarrow&\quad \mu_3 \\ z &= 2 &\leftrightarrow&\quad \lambda \end{aligned}$$

We know the optimal solution in the **primal variables**
$x^* = 3, y^* = 1, z^* = 2, f(x^*, y^*, z^*) = 19$.

What about the **dual variables** μ_i, λ?

We get same solutions to $\mu_1^* = 3, \mu_2^* = 0, \mu_3^* = 3$ because they're not coupled to z direction.
What about λ^*?

Another simple problem with KKT conditions

We use stationarity for the optimal point:

$$0 = \frac{\partial \mathcal{L}}{\partial z} = \frac{\partial f}{\partial z} - \sum_i \lambda_i^* \frac{\partial g_i}{\partial z} - \sum_j \mu_j^* \frac{\partial h_j}{\partial z} = 5 - \lambda^*$$

From which we find:

$$\lambda^* = 5$$

Check interpretation: $\lambda_i = \frac{\partial \mathcal{L}}{\partial c_i}$ with $c_i \to c_i + \varepsilon$.

Find the values of x^*, y^*, μ_i^*

$$\max_{x,y\in\mathbb{R}} f(x,y) = y$$

subject to the following constraints

$$
\begin{aligned}
y + x^2 &\leq 4 \qquad &\leftrightarrow \qquad &\mu_1 \\
y - 3x &\leq 0 \qquad &\leftrightarrow \qquad &\mu_2 \\
-y &\leq 0 \qquad &\leftrightarrow \qquad &\mu_3
\end{aligned}
$$

Simple energy system example

Objective function:

$$\min_{x_1,x_2} f(x_1,x_2) = o_1x_1 + o_2x_2$$

with constraints:

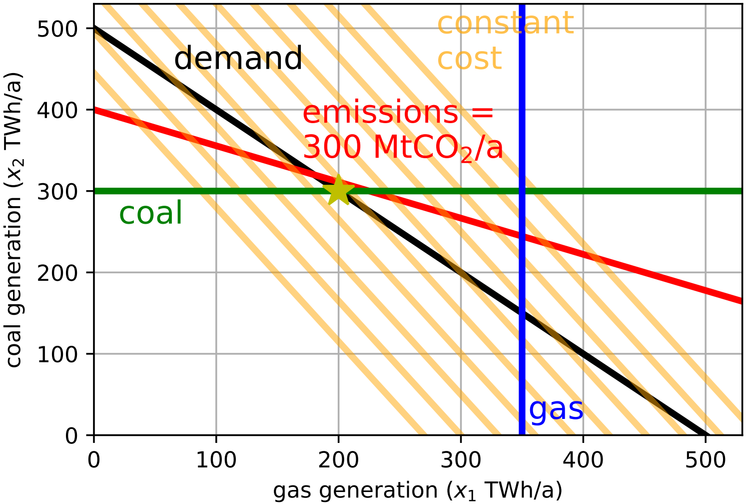
$$\begin{aligned}
x_1 + x_2 &= 500 && \leftrightarrow && \lambda \\
x_1 &\leq 350 && \leftrightarrow && \bar{\mu}_1 \\
-x_1 &\leq 0 && \leftrightarrow && \underline{\mu}_1 \\
x_2 &\leq 300 && \leftrightarrow && \bar{\mu}_2 \\
-x_2 &\leq 0 && \leftrightarrow && \underline{\mu}_2 \\
\frac{e_1}{\eta_1}x_1 + \frac{e_2}{\eta_2}x_2 &\leq 300 && \leftrightarrow && \mu
\end{aligned}$$

Stationarity for $i = 1,2$:

$$0 = \frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial f}{\partial x_i} - \sum_i \lambda_i^* \frac{\partial g_i}{\partial x_i} - \sum_j \mu_j^* \frac{\partial h_j}{\partial x_i} = o_i - \lambda^* - \bar{\mu}_i^* + \underline{\mu}_i^* - \frac{e_i}{\eta_i}\mu^*$$

Simple energy system example

An emissions cap of ≤ 300 MtCO$_2$/a (red) is non-binding.



Simple energy system example

For an emissions limit of 300 we had $x_1^* = 200, x_2^* = 300$. Because so many constraints were non-binding here, we had:

$$\underline{\mu}_1^* = \underline{\mu}_2^* = \bar{\mu}_1^* = \mu^* = 0$$

The only non-zero dual variables are for the binding demand constraint λ^* and the binding capacity constraint for coal $\bar{\mu}_2^*$. Thus we get from stationarity:

$$\lambda^* = o_1 \qquad \lambda^* = o_2 - \bar{\mu}_2^*$$

$\lambda^* = o_1 = 33$ €/$\mathrm{MWh_{el}}$ is the price of supplying an extra unit of electricity. It comes from the gas generator, which still has free dispatchable capacity.

$\bar{\mu}_2^* = o_2 - \lambda^* = o_2 - o_1 = (20 - 33)$ €/$\mathrm{MWh_{el}} = -13$ €/$\mathrm{MWh_{el}}$ is the system cost decrease if the capacity of generator 2 (coal) expands by one unit. This allows us to substitute more expensive generation from gas (at 33€/$\mathrm{MWh_{el}}$) with cheaper generation from coal (at 20€/$\mathrm{MWh_{el}}$).

For a binding emissions limit of 250 we had $x_1^* = 300, x_2^* = 200$. Because none of the generation constraints are non-binding here, we have $\underline{\mu}_1^* = \underline{\mu}_2^* = \bar{\mu}_1^* = \bar{\mu}_2^* = 0$.

The only non-zero dual variables are for the binding demand constraint λ^* and the binding emissions constraint μ^*. Thus we get from stationarity:

$$\lambda^* = o_1 - \frac{e_1}{\eta_1}\mu^* \qquad \lambda^* = o_2 - \frac{e_2}{\eta_2}\mu^*$$

Again we have two equations for two unknowns (λ^* and μ^*) and solve to find:

$$\mu^* = \frac{o_1 - o_2}{\frac{e_1}{\eta_1} - \frac{e_2}{\eta_2}} = \frac{33.3 - 20}{0.333 - 0.75} = -32 \text{ €}/\mathrm{tCO_2}$$

and $\lambda^* = 44$ €$/\mathrm{MWh_{el}}$. How do we interpret these shadow prices?

Looking at the equation for μ^*, $o_1 - o_2$ is the cost in €$/\mathrm{MWh_{el}}$ of replacing coal with gas to reduce emissions, divided by the emissions reduced $\frac{e_2}{\eta_2} - \frac{e_1}{\eta_1}$ in $\mathrm{tCO_2}/\mathrm{MWh_{el}}$, multiplied by -1.

The magnitude of μ^* is the **marginal abatement cost** for CO_2 reduction, i.e. what it costs to reduce the next tCO_2 from the system (e.g. by replacing coal generation by gas). It is negative in this case because μ^* by definition measures the cost reduction by increasing the CO_2 limit by ε, rather than the cost increase by tightening. (Whereas in maximisation problems we have dual feasibility $\mu_j^* \geq 0$, for minimisation problems we have $\mu_j^* \leq 0$ since we obtain a lower objective function when we relax the constraint.)

The market price $\lambda^* = 44$ €/MWh_{el} is higher than either generator's marginal cost. Why?

Because we have to obey the emissions constraint. If we raise demand by 1 MWh_{el}, we increase gas generation, but this increases also CO_2 emissions. To compensate and reduce emissions, we have to substitute some coal for gas. Coal is cheaper than gas, so this also raises the costs.

In another interpretation: we have

$$\lambda^* = o_i - \frac{e_i}{\eta_i}\mu^*$$

so μ^* is adding an effective **carbon price** to the fuel-based marginal cost, thus increasing the marginal cost (remember μ^* is negative).

Optimisation: Solution Algorithms

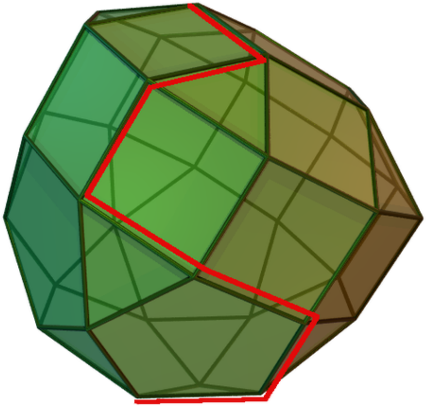
Optimisation solution algorithms

In general finding the solution to optimisation problems is hard, at worst *NP*-hard. Non-linear, non-convex and/or discrete (i.e. some variables can only take discrete values) problems are particularly troublesome.

There is specialised software for solving particular classes of problems (linear, quadratic, discrete etc.).

Since we will mostly focus on linear problems, the main two algorithms of relevance are:

- The **simplex algorithm**
- The **interior-point algorithm**



The simplex algorithm works for linear problems by building the feasible space, which is a multi-dimensional polyhedron, and searching its surface for the solution.

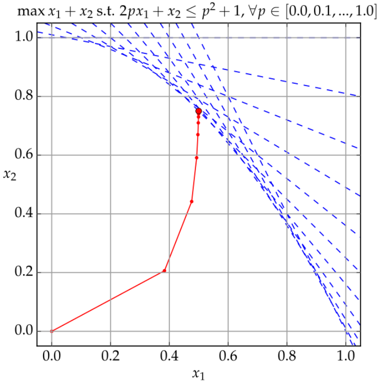
If the problem has a solution, the optimum can be assumed to always occur at (at least) one of the vertices of the polyhedron. There is a finite number of vertices.

The algorithm starts at a feasible vertex. If it's not the optimum, the objective function will increase along one of the edges leading away from the vertex. Follow that edge to the next vertex. Repeat until the optimum is found.

Complexity: On *average* over given set of problems can solve in polynomial time, but worst cases can always be found with exponential time.

Source: Wikipedia

Interior point methods



Interior point methods can be used on more general non-linear problems. They search the interior of the feasible space rather than its surface. They achieve this by extremising the objective function plus a **barrier term** that penalises solutions that come close to the boundary.
As the penality becomes less severe the algorithm converges to the optimum point at the boundary.

Complexity: For linear problems, Karmakar's version of the interior point method can run in polynomial time.

Source: **Wikipedia**

Interior point methods: Barrier method

Take a problem

$$\min_{\{x_\ell, \ell=1,\dots k\}} f(x)$$

such that for

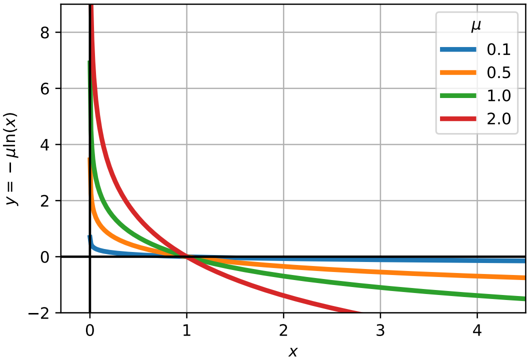
$$g_i(x) = 0 \leftrightarrow \lambda_i, i = 1 \dots n$$
$$x \geq 0$$

Any optimisation problem can be brought into this form. Introduce the **barrier function**

$$B(x, \mu) = f(x) - \mu \sum_{\ell=1}^{k} \ln(x_\ell)$$

where μ is the small and positive **barrier parameter** (a scalar). Note that the barrier term penalises solutions when x comes close to 0 by becoming large and positive.

Barrier term $-\mu ln(x)$ penalises the minimisation the closer we get to $x = 0$. As μ gets smaller it converges on being a near-vertical function at $x = 0$.



Interior point methods: Barrier method: 1-d example

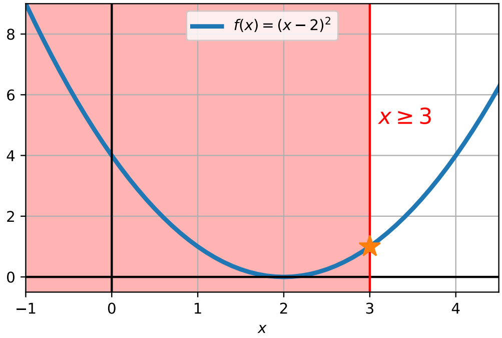
Return to our old 1-d example. We minimise a function of one variable $x \in \mathbb{R}$

$$\min_x (x-2)^2$$

subject to a constraint

$$x \geq 3$$

Solution: $x^* = 3$.

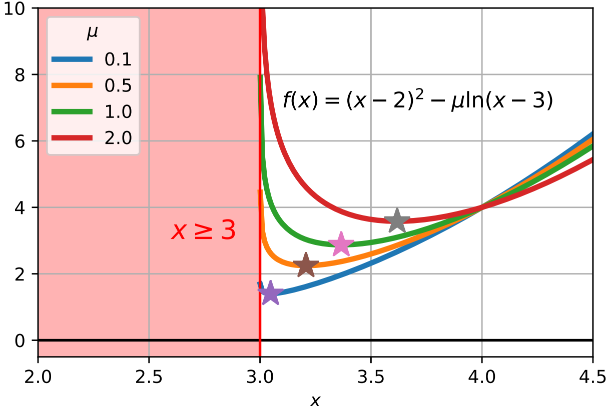


Interior point methods: Barrier method: 1-d example

Now instead minimise the barrier problem without any constraint:

$$\min_x B(x,\mu) = (x-2)^2 - \mu \ln(x-3)$$

Solve $\frac{\partial B(x,\mu)}{\partial x} = 2(x-2) - \frac{\mu}{x-3} = 0$, i.e. at $x^* = 2.5 + 0.5\sqrt{1+2\mu} \to 3$ as $\mu \to 0$.



Interior point methods: Barrier method

Technische Universität Berlin

The problem

$$\min_{\{x_\ell, \ell=1,\dots k\}} \left[f(x) - \mu \sum_{\ell=1}^{k} \ln(x_\ell) \right]$$

such that

$$c_i(x) = 0 \leftrightarrow \lambda_i, i = 1 \dots n$$

can now be solved using the extremisation of the Lagrangian like we did for KKT sufficiency.

Solve the following equation system iteratively using the Newton method to find the x_ℓ and λ_i:

$$\begin{aligned} \nabla_\ell f(x) - \mu \frac{1}{x_\ell} + \sum_i \lambda_i \nabla_\ell c_i(x) &= 0 \\ c_i(x) &= 0 \end{aligned}$$

See this **nice video** for more details and visuals.