# MAT4120

Possible topics and questions on the exam

August Femtehjell

December 6, 2025

## Contents




**Abstract**

This document contains my notes on possible topics and questions for the exam in MAT4120–Mathematical Optimization, taught at the University of Oslo in the spring of 2025. The code for everything, as well as this document, can be found at my GitHub repository: https://github.com/augustfe/MAT4120.

# An introduction to convexity

## Chap. 1

**Exercise 1.1.** What is a convex set? What is an affine set? Examples: the unit cube, hyperplanes.

**Solution 1.1.** A convex set is intuitively a set where for each pair of two points, all the points one the line between those two points are also in the set. This for instance means that the unit circle is convex, while a typical drawing of a star is not, as connecting two "points" places you outside the star. Explicitly, we have that a set $C$ is convex if and only if

$$\lambda x + (1-\lambda)y \in C \tag{1.1}$$

for all $x, y \in C$ and $\lambda \in [0,1]$.

An affine set is practically a linear subspace centred around some point other than zero. It is an affine set if and only if it is the solution set to some linear equation. That is, for a non-empty set $S \subseteq \mathbb{R}^n$, there is a matrix $A \in \mathbb{R}^{m,n}$ and a vector $b \in \mathbb{R}^m$ such that

$$S = \{x \in \mathbb{R}^n : Ax = b\}. \tag{1.2}$$

It can equivalently be written as $C = L + x_0 = \{x + x_0 : x \in L\}$ for some linear subspace $L$ of $\mathbb{R}^n$.

To see the equivalence, let $x_0$ be a solution of the linear equation, such that $Ax_0 = b$. Then, $L$ is the kernel of $A$, as then for $z \in S$ we have

$$Az = A(x + x_0) = Ax + Ax_0 = b \implies Ax = 0. \tag{1.3}$$

**Exercise 1.2.** Is the sum of convex sets also convex?

**Solution 1.2.** Yes. Consider two convex sets $A, B \subseteq \mathbb{R}^n$. We then have that the sum of the sets is given by

$$A + B = \{a + b : a \in A, b \in B\}. \tag{1.4}$$

As $A$ and $B$ are convex, we have for $a_1, a_2 \in A$, $b_1, b_2 \in B$ and $\lambda \in [0,1]$ that

$$\lambda a_1 + (1-\lambda)a_2 + \lambda b_2 + (1-\lambda)b_2 = \lambda(a_1 + b_1) + (1-\lambda)(a_2 + b_2) \in A + B. \tag{1.5}$$

**Exercise 1.3.** What is a cone? Example: First octant in $\mathbb{R}^3$.

**Solution 1.3.** A cone $C$ is a set with the property that a non-negative combination of two points in the cone, is also in the cone. That is, for $\lambda_1, \lambda_2 \geq 0$ and $x, y \in C$, we have

$$\lambda_1 x + \lambda_2 y \in C. \tag{1.6}$$

The first octant in $\mathbb{R}^3$, often denoted $\mathbb{R}^3_+$, is an example of such a space.

**Exercise 1.4.** What is a polyhedron?

**Solution 1.4.** A polyhedron is the solution set to any linear equation $Ax \leq b$.

## Chap. 2

**Exercise 1.5.** What is the convex/affine hull of a set? Is the convex hull a convex set?

**Solution 1.5.** The convex hull $C$ of a set $A$ is the set of all convex combinations of points in $A$. It is also the smallest convex set in which $A$ is a subset.

The affine hull of a set $A$ is the set of all affine combinations of elements in $A$, where an affine combination is given by

$$\sum_{i=0}^{k} \mu_i x_i \tag{1.7}$$

for $\{x_i\}_{i=0}^{k} \in A$ subject to $\sum_{i=0}^{k} \mu_i = 1$ for some $k > 0$. It is also the smallest affine set containing $A$.

**Exercise 1.6.** What is a polytope? Explain why a polytope is compact.

**Solution 1.6.** A polytope is the convex hull of a finite number of points, that is $P = \text{conv}\{x_1, \dots, x_t\} \subset \mathbb{R}^n$. Polytopes are compact as there exists a continuous function $f$ from the standard simplex $S_t$ to $\mathbb{R}^n$, combined with the fact that the standard simplices are compact. A point in $S_t$ can be written as $(\lambda_1, \dots, \lambda_t)$ with $\sum_{i=1}^{t} \lambda_i = 1$ and $\lambda \geq 0$. $f$ then maps that point to $\sum_{i=1}^{t} \lambda_i x_i$.

**Exercise 1.7.** What is affine independence, and what is the connection to linear independence?

**Solution 1.7.** Affine independence is closely related to linear independence. A set of vectors $x_1, \dots, x_t$ are affinely independent, if $\sum_{i=0}^{t} \lambda_i x_i = 0$ and $\sum_{i=0}^{t} \lambda_i = 0$ implies that $\lambda_1 = \cdots = \lambda_t = 0$. Note that linear independence differs, as then the first condition alone implies that all the coefficients are zero. Additionally, if the vectors are affinely independent, then the vectors $x_i - x_1$ for $i = 2, \dots, t$ are linearly independent.

**Exercise 1.8.** What is the (affine) dimension of a set.

**Solution 1.8.** The affine dimension of a set is the maximal number of affinely independent vectors, minus one. It therefore coincides with the usual notion of dimensionality, i.e., the maximal number of linearly independent vectors.

**Exercise 1.9.** What is a simplex?

**Solution 1.9.** A simplex is the convex hull of a set of affinely independent vectors. The standard simplex is for instance given by $S_t = \text{conv}\{e_1, \ldots, e_t\}$, or equivalently $S_t = \{x \in \mathbb{R}^n : x \geq 0, \sum_{i=1}^t x_i = 1\}$.

**Exercise 1.10.** What is the relative topology? Look at the unit disk in $\mathbb{R}^2$ embedded in $\mathbb{R}^3$. What is the relative interior? The interior? The relative boundary?

**Solution 1.10.** Relative topology is a concept which is useful when classifying sets in $\mathbb{R}^n$ with dimensionality less than $n$, by considering the affine hull of that set. For instance, when considering the unit disk in $\mathbb{R}^2$ embedded in $\mathbb{R}^3$, that is,

$$S = \{(x_1, x_2, 0) : x_1^2 + x_2^2 \leq 1\}, \tag{1.8}$$

we find that the interior of $S$ is empty, as any slight change in the $x_3$ coordinate places us outside the set. Explicitly, any ball around a point $x$ in $S$ with radius $r > 0$ contains points outside of $S$. This is obviously not entirely useful.

Because of this, we consider the relative topology, where the relative interior is defined as the intersection of the ball described previously, along with the affine hull of $S$. In this case, $\text{aff}(S)$ is given by the $x_1$-$x_2$ plane. This leads to the intuitive definition of an interior, where for instance $(0, 0, 0)$ is "inside" of $S$, as the intersection restricts us from considering movements in $x_3$. That is,

$$\text{rint}(S) = \{(x_1, x_2, 0) : x_1^2 + x_2^2 < 1\}. \tag{1.9}$$

The relative boundary is defined as usual, given by $\text{rbd}(S) = \text{cl}(S) \setminus \text{rint}(S)$, here given by

$$\text{rint}(S) = \{(x_1, x_2, 0) : x_1^2 + x_2^2 = 1\}, \tag{1.10}$$

as $S = \text{cl}(S)$. Note that as the interior of $S$ is empty, the usual definition of a boundary gives us that $\text{bd}(S) = S$, each point is at the boundary.

**Exercise 1.11.** Can all points in a convex set $S$ be expressed in terms of affinely independent points in $S$? If so, how many do you need? And do the convex combinations of these generate all of $S$?

**Solution 1.11.** Yes, by Caratheodory's theorem, for $S \subseteq \mathbb{R}^n$ we need at most $n+1$ affinely independent points in $S = \text{conv}(S)$. Convex combinations of these selected points do not in general generate all of $S$, and need to be selected for each $x \in S$.

## Chap. 3

**Exercise 1.12.** What is a supporting/separating hyperplane?

**Solution 1.12.** Firstly, a hyperplane is a set of the form

$$H_{\alpha,a} = \{x \in \mathbb{R}^n : a^T x = \alpha\}. \tag{1.11}$$

Accompanying this, we have the halfspaces

$$\begin{aligned} H^+_{\alpha,a} &= \{x \in \mathbb{R}^n : a^T x \geq \alpha\}, \\ H^-_{\alpha,a} &= \{x \in \mathbb{R}^n : a^T x \leq \alpha\}. \end{aligned} \tag{1.12}$$

If a set $S$ is contained entirely in one of these halfspaces, and the intersection is non-zero, we say that that halfspace is a supporting hyperplane. For instance, with a circle, the tangent at a point defines a hyperplane, wherein the circle is entirely on one side.

Given two sets $S$ and $T$, $H$ is a separating hyperplane if $S \subseteq H^{\pm}$ and $T \subseteq H^{\mp}$, that is they are entirely separated by the hyperplane.

**Exercise 1.13.** Suppose that $C$ is closed and convex, and $x$ is not in $C$. Can $x$ and $C$ be separated with a hyperplane?

**Solution 1.13.** As $x \notin C$, and $C$ is closed, there is a ball about $x$ entirely outside of $C$, with some radius $r > 0$. Clearly then, any separating hyperplane is also strongly separating. Let $H$ be the hyperplane supporting $C$ at $p_C(x)$ with normal vector $x - p_C(x)$. As $x \notin C$, we have that $x \neq p_C(x)$. $H$ then separates $x$ and $C$.

**Exercise 1.14.** What does Farkas' Lemma say? Can you sum up its proof? (Separating hyperplane theorem.)

**Solution 1.14.** Farkas' Lemma states that for $A \in \mathbb{R}^{m,n}$ and $b \in \mathbb{R}^n$, there exists an $x \geq 0$ satisfying $Ax = b$ if and only if for each $y \in \mathbb{R}^m$ with $y^T A \geq 0$ it also holds that $y^T b \geq 0$.

The proof considers the convex cone generated by the columns of $A$, i.e., $C = \text{cone}\{a^1, \ldots, a^n\}$. A solution $x \geq 0$ to $Ax = b$ then corresponds to $b$ lying in the cone. The proof then assumes that such a solution exists.

Further, if $y^T A \geq 0$, then

$$y^T b = y^T (Ax) = (y^T A)x \geq 0, \tag{1.13}$$

as both $x$ and $y^T A$ are assumed non-negative.

Note now that $C$ is closed and convex. If $Ax = b$ has no non-negative solution, then $b \notin C$, such that $b$ and $C$ can be strongly separated by the separating hyperplane theorem. There is therefore a non-zero vector $y \in \mathbb{R}^m$ and $\alpha > 0$ such that $y^T x \geq \alpha$ for all $x \in C$ and $y^T b < \alpha$. As $C$ is a cone, $0 \in C$, such that $\alpha \leq 0$.

Furthermore, we must have that $y^T x \geq 0$. Suppose for contradiction that we had some $x_0 \in C$ such that $y^T x_0 < 0$. We would then have

$$\alpha \leq y^T x_0 < 0. \tag{1.14}$$

As $C$ is a cone, we also have that $\lambda x_0 \in C$ for all $\lambda > 0$. We could then choose $\lambda$ sufficiently large such that $\lambda y^T x_0 < \alpha$, a contradiction.

Therefore, $y^T a^j \geq 0$, as each $a^j \in C$, so $y^T A \geq 0$. Furthermore, we in this case have that $y^T b < \alpha \leq 0$, proving the other direction.

## Chap. 4

**Exercise 1.15.** What is a face of a convex set? What are the faces of different dimensions for the unit cube of $\mathbb{R}^3$?

**Solution 1.15.** A face $F$ of a convex set $C$ is a set wherein $x = \lambda x_1 + (1-\lambda)x_2 \in F$ implies that both $x_1$ and $x_2$ lie in $F$, for each such possible convex combination in $C$.

The faces of dimension 0 for the unit cube of $\mathbb{R}^3$ is each vertex. Of dimension 1, we have the edges, while we for dimension 2 have the surfaces typically called faces. For dimension 3, we have the full cube. Not sure if defined, but the face of dimension $-1$ might logically be concluded to be the empty set, as it doesn't contain any affinely independent points, and fulfils the requirements for a face.

**Exercise 1.16.** What is an extreme point? What are those for the unit cube and the unit sphere?

**Solution 1.16.** An extreme point of a set $C$ is a point $x \in C$ where $x_1, x_2 \in C$ such that $x = \frac{1}{2}(x_1 + x_2)$ implies that $x_1 = x = x_2$.

For the unit cube, these are the vertices, while on the unit sphere this is every point on the boundary, satisfying $\|x\| = 1$.

**Exercise 1.17.** Given a polytope $\text{conv}(x_1, \ldots, x_t)$, what can you say about the extreme points?

**Solution 1.17.** We have that the only candidates for extreme points are $x_1, \ldots, x_t$, however we do not necessarily have that all $x_1, \ldots, x_t$ are extreme.

**Exercise 1.18.** Let $S$ be a subset of vectors with all components being either 0 or 1. Explain why the extreme points of $\text{conv}(S)$ are precisely $S$.

**Solution 1.18.** We have that $S$ is a finite set of vectors, so the only candidates for the extreme points are precisely those vectors. In order for one of the points to not be an extreme point, we require that we write it as a combination of other points. Considering a specific component, say a 0, we would either need for both corresponding components to be zero, or have opposite sign. Clearly, as there are no negative components involved, they must be equal. Similarly, for a component 1, we must have either numbers above and below, or equivalence. As none are above, they must also be equal. The vectors are therefore identical, so the points are extreme.

**Exercise 1.19.** What is the recession cone of a closed convex set?

**Solution 1.19.** The recession cone of a closed convex set $C$ contains all direction vectors, such that each halfline starting at a point in $C$ is entirely contained in $C$. Explicitly, the set of halflines starting at a point $x \in C$, given by

$$\operatorname{rec}(C, x) = \{z \in \mathbb{R}^n : x + \lambda z \in C \ \forall \lambda \geq 0\}, \tag{1.15}$$

is actually independent of the chosen point $x$. Read up on why it is independent.

**Exercise 1.20.** Suppose $C$ is a convex compact set. What does Minkowski's theorem say about $C$? ($C = \operatorname{conv}(\operatorname{ext}(C))$.)

**Solution 1.20.** Minkowski's theorem tells us that a convex compact set $C$ is the convex hull of its extreme points. This build upon the inner description of closed convex sets, which states that a non-empty and line-free closed convex set $C$ is the convex hull of its extreme points *and extreme half-lines*. As $C$ in our case is compact, it is necessarily bounded, and therefore contains no half-lines.

**Exercise 1.21.** What is a vertex of a polyhedron? What is the connection between vertices and extreme points?

**Solution 1.21.** We consider a non-empty polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ for some $A \in \mathbb{R}^{m,n}$ and $b \in \mathbb{R}^m$. Each vertex in a polyhedron is then the solution of $n$ linearly independent equations from the linear system.

As each such vertex lies in the polyhedron, $x_0 \in P$, it is also an extreme point. Assume for contradiction that we can write $x_0 = \frac{1}{2}(x_1 + x_2)$ with $x_1 \neq x_2 \in P$. Then,

$$A_0 x_0 = A_0(\tfrac{1}{2}(x_1 + x_2)) = \tfrac{1}{2} A_0 x_1 + \tfrac{1}{2} A_0 x_2 < b, \tag{1.16}$$

where we have a strict inequality as $x_0$ is a unique solution, as the equations are linearly independent, contradicting $A_0 x_0 = b$.

**Exercise 1.22.** Why is the intersection of two polytopes again a polytope?

**Solution 1.22.** Consider two polytopes

$$\begin{aligned} P_1 &= \{x : A_1 x \leq b_1\}, \\ P_2 &= \{x : A_2 x \leq b_2\}. \end{aligned} \tag{1.17}$$

The intersection is then given by

$$P_1 \cap P_2 = \{x : A_1 x \leq b_1 \text{ and } A_2 x \leq b_2\}, \tag{1.18}$$

which we can achieve as a single polygon with

$$P_1 \cap P_2 = \left\{ x : \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x \leq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\}. \tag{1.19}$$

**Exercise 1.23.** What does the main theorem for polyhedra say?

**Solution 1.23.** The main theorem for polyhedra states that for each polyhedra $P \subseteq \mathbb{R}^n$ there are finite sets $V$ and $Z$ such that

$$P = \operatorname{conv}(V) + \operatorname{cone}(Z). \tag{1.20}$$

Additionally, for finite sets $V$ and $Z$, $\operatorname{conv}(V) + \operatorname{cone}(Z)$ is a polyhedron. For a pointed polyhedra, $V$ is the set of vertices, while $Z$ is the set of direction vectors for each extreme half-line.

## Chap. 5

**Exercise 1.24.** What is a convex function?

**Solution 1.24.** A convex function is a function $f : C \to \mathbb{R}$, where $C \subseteq \mathbb{R}^n$ is convex, satisfying

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{1.21}$$

for all $x, y \in A$ and $\lambda \in [0,1]$.

**Exercise 1.25.** What is the slope function, and its connection to convexity?

**Solution 1.25.** We here consider points $P_x = (x, f(x))$, and define the slope by

$$\operatorname{slope}(P_x, P_y) = \frac{f(x) - f(y)}{x - y}, \tag{1.22}$$

for $x < y$. Adjacent to this is the slope function, defined by

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}, \tag{1.23}$$

for some point $x_0 \in C$. The function $f$ is then convex if and only if the slope function is increasing on $\mathbb{R} \setminus \{x_0\}$.

**Exercise 1.26.** What is the epigraph of a function? What can you say about the epigraph of a convex function?

**Solution 1.26.** The epigraph of a function is the set defined by

$$\text{epi}(f) = \{(x, y) \in \mathbb{R}^{n+1} : y \geq f(x)\ \forall x \in C\}. \tag{1.24}$$

That is, the set "above" the function. A function $f$ is convex if and only if epi($f$) is a convex set.

## Chap. 6

**Exercise 1.27.** Write down the general form of an optimisation problem with equality and inequality constraints. What do we mean by regular points and active constraints?

**Solution 1.27.** The general form of an optimisation problem with equality and inequality constraints is given by

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & \\ & h_i(x) = 0 \quad \text{for } i = 1, \ldots, m \\ & g_j(x) = 0 \quad \text{for } j = 1, \ldots, r. \end{aligned} \tag{1.25}$$

Equivalently, we write $h(x) = 0 \in \mathbb{R}^m$ and $g(x) \leq 0 \in \mathbb{R}^r$. Active constraints are those which are active in an optimal solution, such that we only need to consider equalities, being $h$ and the inequalities which are active ($g_j(x) = 0$).

A feasible point $x^*$ is called regular if the gradients of $h_i$ and the active $g_j$ are linearly independent at $x^*$.

**Exercise 1.28.** What do we mean by the KKT conditions? What is the connection with Lagrange's multiplier method?

**Solution 1.28.** The KKT conditions are optimality conditions for optimisation problems with equality and inequality constraints. Along with the conditions, we consider the Lagrangian

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{r} \mu_j g_j(x). \tag{1.26}$$

A triple $(x^*, \lambda^*, \mu^*)$ is said to satisfy the KKT conditions if we have primal feasibility, i.e.

$$g(x^*) \leq 0 \quad \text{and} \quad h(x^*) = 0, \tag{1.27}$$

dual feasibility,

$$\mu_j \geq 0 \quad \text{for } j = 1, \ldots, r, \tag{1.28}$$

stationarity,

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{r} \mu_j \nabla g_j(x^*) \tag{1.29}$$

and complimentary slackness, $\mu_i^* g_i(x^*) = 0$. If $x^*$ is regular, then the multiplies are unique, as the active inequality constraints are linearly independent.

# Notes on combinatorial optimisation

## Chap. 1

**Exercise 2.1.** State the general combinatorial optimisation problem. What is the node-edge incidence matrix of graph?

**Solution 2.1.** In a combinatorial optimisation problem, we consider a class $\mathcal{F}$ which is a subset of some finite ground set $E$. Associated with the problem is a weight function $w : E \to \mathbb{R}$, defined for an $F \in \mathcal{F}$ as $w(F) = \sum_{e \in F} w(e)$. The problem is then simply given by

$$\max\{w(F) : F \in \mathcal{F}\}. \tag{2.1}$$

For a graph $G = (V, E)$, the node-edge incidence matrix is the matrix $A \in \{0, 1\}^{V \times E}$, such that the column corresponding to an edge $e = (u, v)$ has two non-zero entries, corresponding to the nodes $u$ and $v$.

**Exercise 2.2.** What is the forest polytope $(F(G))$ of a graph?

**Solution 2.2.** The forest polytope is the convex hull of all incidence vectors of all forests in $G$, where an incidence vector is the vector $\chi^F$ with $\chi^F_e = 1$ if $e \in F$ and zero otherwise.

**Exercise 2.3.** Can you describe $F(G)$ as a polyhedron in $(Q)$? What are the constraints? $(x(E[S]) \geq |S| - 1)$. Say something about the proof (look at the vertices, characterise them as optimal solutions of an LP problem, show by complementary slackness that the greedy algorithm finds optimal solutions. This is an incidence vector).

**Solution 2.3.** There is a lemma which states that $x \in \{0, 1\}^E$ is an incidence vector of a forest if and only if

$$x(E[S]) \leq |S| - 1 \qquad \text{for all } S \subseteq V. \tag{2.2}$$

Considering the $Q$ be the polytope defined by

$$\begin{aligned} x_e &\geq 0 && \text{for all } e \in E \\ x(E[S]) &\leq |S| - 1 && \text{for all } S \subseteq V. \end{aligned} \tag{2.3}$$

The integral vectors in $Q$ are precisely the incidence vectors by the lemma, so $F(G) \subseteq Q$.

Let now $\bar{x}$ be a vertex of $Q$, meaning we can find it as a unique optimal solution $\bar{x} = \arg\max_{x\in Q} c^T$, for some suitable $c$. The dual of this LP problem is given by

$$\begin{aligned} \text{minimize} \quad & \sum_{S\subseteq V} y_S(|S|-1) \\ \text{subject to} \quad & \\ & \sum_{\{S:e_r\in E[S]\}} y_S \geq c_e \qquad \text{for all } e\in E \\ & y_S \geq 0 \qquad \text{for all } S\subseteq V. \end{aligned} \tag{2.4}$$

Applying the greedy algorithm to the primal problem, we get a solution $F = \{e_1, \ldots, e_s\}$, and assume it is in that order. At the $i$-th iteration, we therefore find $e_i = (u, v)$, joining together two components in the current solution into the new component $V_i$.

For the dual problem, we need to define the dual solution $y$. We let $y_S = 0$ for all $S \notin \{V_1, \ldots, V_s\}$. We define the values for the remaining solutions recursively, starting from the back with $y_{V_s} = c(e_r)$. By the complementary slackness condition for an edge $e$, as $x'_e > 0$, we have that $\sum_{\{S:e_r\in E[S]\}} y_S = c_e$. By our definition of $y_{V_s}$, we are then required to set $y_S = 0$ for all sets not including the endpoints of $e_s$.

We define the remaining components of $y$ with the index set

$$I(j) = \{i : j+1 \leq i \leq s \text{ and both end nodes of } e_j \text{ are in } V_i\}, \tag{2.5}$$

for $j \in \{1, \ldots, r-1\}$. We then set

$$y_{V_j} = c(e_j) - \sum_{i\in I(j)} y_{V_i} \tag{2.6}$$

for $j = r-1, r-2, \ldots, 1$.

We then check that the solution $y$ is dual feasible, $x'$ is primal feasible and that the complementary slackness condition holds. Both solutions are then optimal in their respective problems, and so by uniqueness we have $\bar{x} = x'$, showing that every vertex of $Q$ is integral, so $Q = F(G)$.

**Exercise 2.4.** What is a formulation for a set $S$?

**Solution 2.4.** With $S \subseteq \{0,1\}^n$, a formulation for $S$ is a polyhedron $P \subseteq \mathbb{R}^n$ which satisfies $P \cap \{0,1\}^n = S$.

**Exercise 2.5.** What is a Hamiltonian tour in a graph? Can you state (in)equalities which characterise incidence vectors of Hamiltonian cycles? (Subtour/degree constraints)

**Solution 2.5.** A Hamiltonian tour is a tour which visits each vertex in a graph exactly once. A Hamiltonian cycle similarly requires that the tour starts and stops in the same node. We can formulate this with inequalities, by requiring that each strict subset of the graph contains at most $|S|-1$ edges, such that no interior cycle exists. Additionally, each vertex must have degree two.

**Exercise 2.6.** What is the travelling salesman problem?

**Solution 2.6.** TSP is the problem of finding a Hamiltonian cycle which minimizes some distance metric, in the simplest case the physical distance a salesman must travel to pass through as number of cities, before ending up back at home.

**Exercise 2.7.** What is a separation oracle? Why do we use these?

**Solution 2.7.** A separation oracle is an algorithm, which from a trial solution $\bar{x}$ to some problem defined on a polyhedron

$$P = \{x \in \mathbb{R}^n : Ax \leq b\}, \tag{2.7}$$

either confirms that $\bar{x} \in P$, or returns a constraint which $\bar{x}$ violates.

We use these as we can very quickly encounter problems with a large number of constraints, wherein we do not need to actually consider all in order to get a solution. Some of the inequalities might be able to be reduced to each other. Figuring this out however can be a lot of work, so just getting the violated ones can be a nice time save.

## Chap. 2

**Exercise 2.8.** What is an integral polyhedron? What is a rational polyhedron?

**Solution 2.8.** From a polyhedron $P \subset \mathbb{R}^n$ we define the integer hull $P_I$ by

$$P_I = \operatorname{conv}(P \cap \mathbb{Z}^n). \tag{2.8}$$

A polyhedron $P$ is then integral if $P = P_I$. A polyhedron is rational if it is defined by linear inequalities with only rational numbers.

**Exercise 2.9.** What is a totally unimodular matrix? Is a matrix with $-1$ in all entries TU?

**Solution 2.9.** A totally unimodular matrix has the property that each square submatrix has determinant $-1$, 0 or 1. A matrix with all negative ones is TU, as each submatrix contains only linearly dependent rows, and thus all have determinant 0, except the trivial submatrices which only have the element $-1$.

**Exercise 2.10.** If $A$ is TU and $b$ is integral, what can you say about the polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$?

**Solution 2.10.** We then have that the polyhedron is integral. The proof is based around Cramer's rule, from which we can show that if a matrix $C$ is integral with $\det C$ is either $-1$ or 1, then $C^{-1}$ is also integral.

**Exercise 2.11.** When is the node-edge incidence matrix of a graph TU?

**Solution 2.11.** The node-edge incidence matrix $A_G$ of a graph $G$ is TU if and only if $G$ is bipartite. The proof here is based on the Ghouila–Houri criterion, which states that a $-1, 0, 1$-matrix $A$ is TU if and only if for each subset of rows $J \subseteq \{1, \ldots, n\}$ there are partitions $J_1, J - 2$ of $J$ such that

$$\left| \sum_{j \in J_1} a_{ij} - \sum_{j \in J_2} a_{ij} \right| \quad \text{for } i = 1, \ldots, m. \tag{2.9}$$

**Exercise 2.12.** Explain the following concepts in a graph:

- Matching,
- Node cover,
- Node packing,
- Edge cover.

**Solution 2.12.** A matching is an edge subset wherein no node is incident to more than one edge. Each node in the subset is therefore either "matched" with another node, or solitary.

A node cover is a node subset such that each edge has an endnode in the cover. A node packing on the other hand is a node subset wherein no nodes are adjacent. An edge cover is finally an edge subset such that each node is connected to one of the edges in the subset.

In particular, in a bipartite graph we have a minmax relation where the maximum cardinality of a node packing equals the minimum cardinality of an edge cover.

## Chap. 3

**Exercise 2.13.** Explain Christofides algorithm for finding a Hamiltonian tour in a weighted graph. What does it mean that this is a factor 2-approximation algorithm?

**Solution 2.13.** Christofides algorithm is a heuristic algorithm for the metric TSP, wherein the weight function satisfies the triangle inequality, i.e.

$$w(u, v) \leq w(u, z) + w(z, v) \tag{2.10}$$

for all $u, v, z \in V$. It is a 2-approximation algorithm in the sense that the solution we obtain is at most a factor of 2 away, i.e.

$$w(x) \leq \alpha v(Q). \tag{2.11}$$

The algorithm consists of first finding a spanning tree $T$ of $G$. Then, we double every edge of $T$ in order to obtain an Eulerian graph, and then find an Eulerian tour $\mathcal{T}$ in this graph. We then return the Hamiltonian tour that visits the vertices of $G$ in the order of their first appearance in $\mathcal{T}$.

**Exercise 2.14.** What is an enumeration tree? What do we mean by branch and bound?

**Solution 2.14.** An enumeration tree is a way to list out the possible solutions. Consider $S \in \{0, 1\}^n$. We build the tree by fixing the first node, according to some fixed ordering. That is, $S^0$ includes all vectors $x \in S$ with $x_0 = 0$. This set then becomes the root node in the tree.

We build upon this tree with the child nodes $S^{0,0}$ and $S^{0,1}$, which is those $x$ with $x_0 = 0$ and respectively $x_1 = 0$ and $x_1 = 1$. This tree is then built recursively, with the final leaf nodes being complete enumerations of the solutions.

This is not entirely practical however, as the number of nodes in the tree grows exponentially. It is however useful, as we may need only consider a small part of the tree in order to solve the problem. This stems from the fact that we can in some cases prune a node $u$, and therefore the entire branches below $u$, meaning we do not need to consider any of the problems in $Q(u')$ for $u'$ below $u$.

The conditions for pruning a node $u$ is that one of the following holds:

1. (Infeasibility) $S(u)$ is empty.

2. (Optimality) An optimal solution of $Q(u)$ is known.

3. (Value dominance) $v(Q(u)) \leq v(Q)$.

A relative of the enumeration tree is the branching tree, where we do not necessarily always fix node $k$ in layer $k$, but rather may fix $i$ in some part, and $k$ in another.

The branch and bound algorithm is a way to work through enumeration trees. It consists of:

Step 1. (Initialization.) Let $V_n = \{v_r\}$, $z_L = -\infty$ and $z_U = \infty$.

Step 2. (Termination.) If $V_n = \emptyset$, the current best solution $x^*$ is optimal, terminate the algorithm.

Step 3. (Node selection and processing.) Select a node $u \in V_n$ and set $V_n \coloneqq V_n \setminus \{u\}$. Solve the LP relaxation $R(u)$. Let $z(u)$ and $x(u)$ respectively denote the optimal value and solution to $R(u)$.

Step 4. (Pruning.)

*(i)* If $R(u)$ is infeasible, go to Step 2.

*(ii)* If $x(u) \in S(u)$ and $z(u) > z_L$, update the best solution by setting $x^* = s(u)$ and $z_L = z(u)$, go to Step 2.

*(iii)* If $z(u) \leq z_L$, go to Step 2..

*(iv)* Otherwise, continue to Step 5.

Step 5. (Branching.) Add two new nodes $u_0$ and $u_1$ to $V_n$, each being a child of node $u$ such that $S(u_0)$ and $S(u_1)$ is a partition of $S(u)$. Go to Step 2.

**Exercise 2.15.** What do we mean by pruning? Can you state a case where pruning can be applied?

**Solution 2.15.** We prune a node by removing it, and all child nodes, from the enumeration tree. We can apply pruning if for instance the optimal value $v(Q(u))$ is lower (or equal) to $v(Q)$, or e.g. if $S(u)$ is empty.

**Exercise 2.16.** What do we mean by valid inequalities for a set $S$?

**Solution 2.16.** An inequality is considered valid if each point $x \in S$ satisfies it, i.e.

$$S \subseteq \{x \in \mathbb{R}^n : a^T x \leq \alpha\}, \tag{2.12}$$

for a given inequality $a^T x \leq \alpha$.

# Network flows and combinatorial optimisation

## Chap. 1

**Exercise 3.1.** What is a flow in a graph?

**Solution 3.1.** A flow is a function $x : E \to \mathbb{R}$, or equivalently $x \in \mathbb{R}^E$. $x$ therefore assigns a value $x(e)$ to each edge $e \in E$. We typically require it to be non-negative.

**Exercise 3.2.** What is the divergence of a flow?

**Solution 3.2.** The divergence of a flow $x$ is the function $\mathrm{div}_x : V \to \mathbb{R}$, given by

$$\mathrm{div}_x(v) = \sum_{e \in \delta^+(v)} x(e) - \sum_{e \in \delta^-(v)} x(e). \tag{3.1}$$

It is therefore the difference between the outflow and inflow for all nodes connected to $v$.

**Exercise 3.3.** What is a circulation in a graph?

**Solution 3.3.** A circulation in a graph is a flow such that $\mathrm{div}_x(v) = 0$ for all $v \in V$.

**Exercise 3.4.** What does Hoffman's circulation theorem say? The proof is rather technical, but you should understand the role of the auxiliary graph therein.

**Solution 3.4.** Hoffman's circulation theorem states that for functions $l, u : E \to \mathbb{R}$ satisfying $l \leq u$, then there exists a circulation $x$ such that

$$l \leq x \leq u \tag{3.2}$$

if and only if

$$\sum_{e \in \delta^-(S)} l(e) \leq \sum_{e \in \delta^+(S)} u(e) \tag{3.3}$$

for all $S \subseteq V$.

Intuitively, the requirement is that for each subset of nodes, the capacity of what can flow out from the subset must be more than what needs to flow in.

Look closer at auxiliary graphs.

**Exercise 3.5.** What is an $st$-flow, and the value of a flow?

**Solution 3.5.** An $st$-flow is a flow along a path from $s$ to $t$, where the divergence of each internal node in the flow is zero. The value of the flow, how much flows from $s$ to $t$ is then equal to the flow out of $s$ by the flow balance equations.

**Exercise 3.6.** What is an $st$-cut?

**Solution 3.6.** An $st$-cut is an edge subset of the form $K = \delta^+(S)$ for a vertex subset $S \subseteq V$ with $s \in S$ and $t \notin S$.

**Exercise 3.7.** What is the maximum $st$-flow problem, and what is the minimum $st$-cut problem? What is the connection between the two?

**Solution 3.7.** The maximum $st$-flow problem is the problem of finding the $st$-flow which maximizes the value of the flow. Similarly, the minimum $st$-cut problem the is the problem of finding the cut $K$ which minimizes the capacities in the cut.

Through duality theory, the optimal value for both problems is the same.

**Exercise 3.8.** How is Hoffman's theorem used in the proof? Algorithm (Ford Fulkerson and $x$-augmenting paths.)

**Solution 3.8.** The proof uses the theorem by adding the edge $(t, s)$ with $l(e) = u(e) = M$, the minimum cut capacity. We therefore just need to show that there is a circulation which satisfies the given constraints, as then $\text{div}_x(s) = 0$, so $\sum_{e\in\delta^+(s)} x(e) = M$. We set $l(e) = 0$ and $u(e) = c(e)$ for each other edge $e \in E$.

Consider then a subset $S$ with $s \in S$ and $t \notin S$. We then have that

$$\sum_{e\in\delta^-(S)} l(e) = M + 0 = M, \tag{3.4}$$

and

$$\sum_{e\in\delta^+(S)} u(e) = \sum_{e\in\delta^+(S)} c(e) = \text{cap}_c(\delta^+(S)). \tag{3.5}$$

The second condition in Hoffman's circulation theorem then reduces to

$$M \leq \text{cap}_c(\delta^+(S)) \tag{3.6}$$

for all $S \subseteq V$, which we know to be true as $M$ is minimal.

The Ford-Fulkerson algorithm starts out with a zero flow $x = 0$. It then repeatedly looks for an $x$-augmenting path $P$ in $D_x$, and if it exists then increase the forward edges in the path by the determined amount $\varepsilon$, while decreasing the value of the backward edges by $\varepsilon$.

## Chap. 2

**Exercise 3.9.** What is a doubly stochastic matrix? Is the set of doubly stochastic matrices convex?

**Solution 3.9.** A doubly stochastic matrix is a matrix with non-negative entries where each row and column sums to one. The set is convex, as it is the convex hull of the set of permutation matrices.

**Exercise 3.10.** What does it mean that a vector $x$ is majorized by a vector $y$?

**Solution 3.10.** A vector is majorized if

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i \tag{3.7}$$

for all $k = 1, \ldots, n-1$, and $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$.

**Exercise 3.11.** State the Gale Ryser theorem, in particular its connection to majorization and given row/column sums.

**Solution 3.11.** Let $R = (r_1, \ldots, r_m)$ and $S = (s_1, \ldots, s_n)$ be non-negative integral vectors with the same sum. The Gale Ryser theorem then tells us that is an $m \times n$ $(0,1)$-matrix $A$ with

$$\begin{aligned} \sum_{j=1}^{n} a_{ij} &= r_i \qquad (i \leq m) \\ \sum_{i=1}^{m} a_{ij} &= s_j \qquad (j \leq n) \end{aligned} \tag{3.8}$$

if and only if $S \preceq R^*$.

Here, $R^*$ is the conjugate of the integral vector $R$, defined by

$$r_k^* = |\{i : r_i \geq k\}| \qquad (k \leq n). \tag{3.9}$$

Therefore, there exists a $(0,1)$-matrix with row sums $R$ and columns sums $S$ if and only if $S$ majorizes the conjugate of $R$.

**Exercise 3.12.** Let $R$ and $S$ be vectors with positive integers. When can we find an integral matrix $A$ with row sums $R$ and column sums $S$, and so that $0 \leq A \leq C$ (with $C$ integral)? Can this be related to graphs?

**Solution 3.12.** Such a matrix exists if and only if

$$\sum_{\substack{i \in I \\ j \in J}} c_{ij} \geq \sum_{j \in J} s_j - \sum_{i \notin I} r_i, \tag{3.10}$$

for all $I \subseteq \{1, 2, \ldots, m\}$ and $J \subseteq \{1, 2, \ldots, n\}$.

This coincides with an integral flow of a bipartite graph with nodes $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$, with $b(u_i) = r_i$ and $b(v_j) = s_j$. An integral flow $x$ with $\mathrm{div}_x = b$ and $0 \leq x \leq c$ then corresponds to the matrix $C$ with the properties desired.

**Exercise 3.13.** What are the vertices/extreme points for the set of doubly stochastic matrices?

**Solution 3.13.** The vertices are the permutation matrix. They clearly cannot be written as convex combinations of other doubly stochastic matrices. Additionally, for a doubly stochastic matrix $A$, we can trace out a cycle going through non-integer entries. We can then construct a matrix with a positive one in even places, and negative ones in the odd places. As each entry lies in the open interval $(0, 1)$, there is an $\varepsilon > 0$ such that $A \pm \varepsilon V$ is a doubly stochastic matrix. Therefore, no other doubly stochastic matrix is an extreme point.