Introduction to Linear Programming

Anthony Papavasiliou, National Technical University of Athens (NTUA)

Source: appendix A, Papavasiliou [1]

Contents

- Mathematical programming models
- Linear programming problems
- Γραφική επίλυση μοντέλων γραμμικού προγραμματισμού
- Γραμμικά προβλήματα σε κανονική μορφή
- Ακραία σημεία και ο αλγόριθμος simplex
- Δυϊκότητα γραμμικού προγραμματισμού
- Ευαισθησία
- Συνθήκες KKT για γραμμικά προγράμματα
- Αναπαράσταση κατά τμήματα γραμμικών συναρτήσεων
- Ανάλυση ευαισθησίας

Mathematical programming models

Mathematical programming models

- Mathematical programming models are decision-making tools in complex systems
- **Decision variables**: denoted typically as x
- Belong to a **feasible set** X which is a subset of $\mathbb{R}^n$, $x \in X \subseteq \mathbb{R}^n$
- The goal that we wish to optimize can be expressed in the form of a function, with $f: \mathbb{R}^n \to \mathbb{R}$ an **objective function** that scores the performance of decision x
  - Math programs aim at either minimizing or maximizing the objective function f

Minimization problem

- Minimization problem:

$$\min_x f(x)$$
$$\text{s.t. } x \in X$$

- The x under the «min» operator indicates decision variables
- The «s.t.» which precedes $x \in X$ corresponds to the expression «subject to», and indicates that the constraints of the problem follow
    - Sometimes «s.t.» is written out as «subject to», or ommitted

Maximization problem

- Maximization problem:

$$\max_x g(x)$$
$$\text{s.t. } x \in X$$

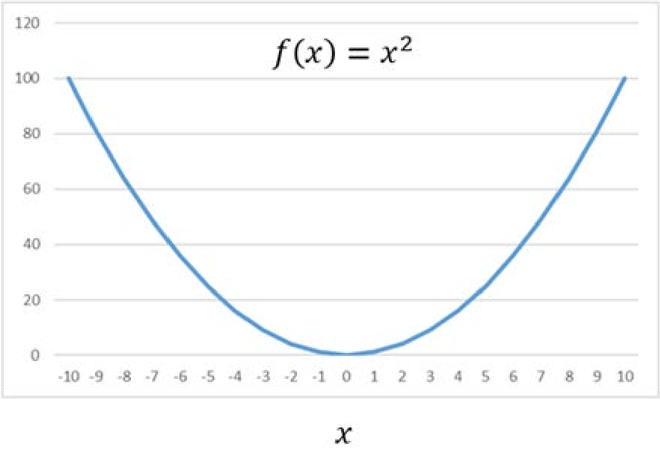
- Equivalent to the minimization problem of the previous slide if $g(x) = -f(x)$

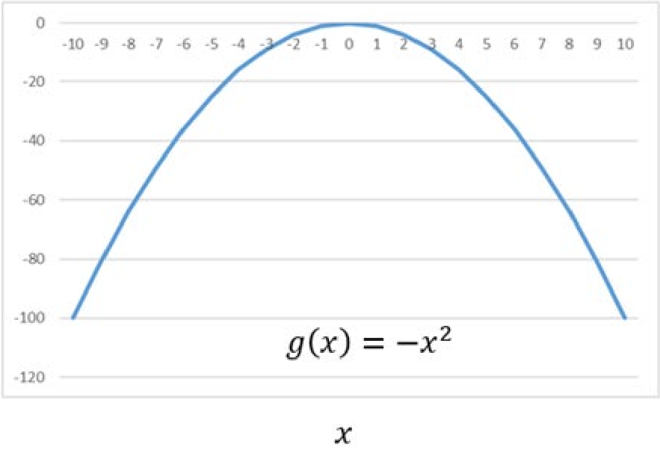
Feasible, infeasible and unbounded problems

- A solution is **feasible** if $x \in X$
- A solution x^* is **optimal** if it is better than all other feasible solutions, which means $f(x^*) \leq f(x)$ for all $x \in X$ in the case of minimization
  - or $f(x^*) \geq f(x)$ for all $x \in X$ in the case of maximization
- A math program is **infeasible** if $X = \emptyset$
- A math program is **unbounded** if the value of its objective function can become arbitrarily low in the space of feasible solutions (in the case of minimization)
  - or arbitrarily high (in the case of maximization)
- Infeasible and unbounded math programs are indications that the underlying problem is not well-defined
- There is no optimal solution for infeasible and unbounded math programs

Example A.1: continuous problem

- Suppose that we minimize the function x^2 on the real numbers
- The objective function is $f(x) = x^2$
- The constraints are $X = \mathbb{R}$
- The math programming problem is expressed as $\min_x x^2$, or equivalently as $\max_x -x^2$





Example A.2: discrete problem

- Consider the decision problem of buying airplane tickets for an upcoming trip from Athens to Brussels
- A roundtrip that does not include the weekend costs \$600
- A roundtrip that includes the weekend costs \$400
- One-way tickets cost \$350 each
- The feasible set can be enumerated, with $X = \{1,2,3\}$
- Objective function: $f(1) = 600$, $f(2) = 400$, $f(3) = 700$
- If the goal is to minimize cost, then $x^* = 2$ and $f(x^*) = 400$

Solution methods

- There is no silver bullet for solving general math programs
- The continuous problem can be solved using differential calculus (derivative equals zero)
- The discrete problem can be solved by enumerating feasible solutions

Problem classes

- None of these strategies (enumeration, or setting the derivative to zero) is fully general or appropriate for large-scale problems
- What determines whether a math program is hard or not?
  - One thought: if the objective function f and constraints X are *linear*, then the problem is easy
- Wrong! The crucial property is not linearity, but *convexity*
  - If the function f is convex, and if the set X is convex, then the problem is «easy»

Linear programming problems

Linear programming problems

- A **linear programming model** is a mathematical programming model where:
  - the objective function f is a linear function of decision variables
  - the constraints X are linear equalities or inequalities of the decision variables

Example A.3: economic dispatch problem

- Consider the economic dispatch problem where two generators produce electricity in order to cover a load of 100 MWh
- Offer generator 1: 60 MWh at 20 €/MWh
- Offer generator 2: 80 MWh at 50 €/MWh
- Goal: cover load at least cost

Economic dispatch model

Decision variables: p_1, p_2 (what do they correspond to?)

$$\min_{p_1,p_2} 20 \cdot p_1 + 50 \cdot p_2$$

$$p_1 + p_2 \geq 100$$

$$p_1 \leq 60$$

$$p_2 \leq 80$$

$$p_1, p_2 \geq 0$$

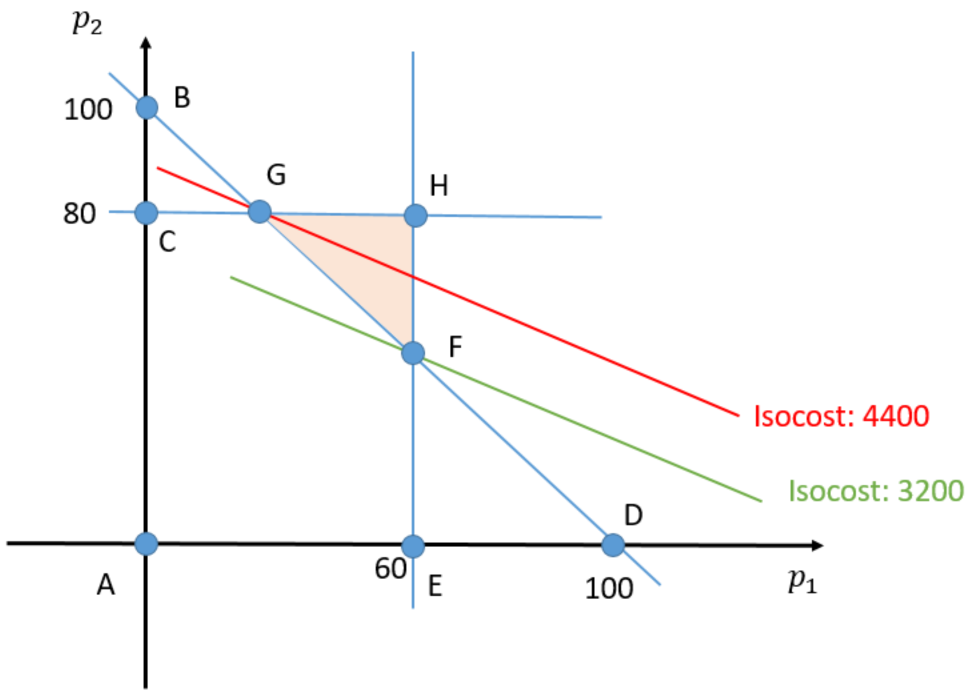
Direction of inequalities and sign of variables

- Linear programming inequalities can be of the ($\leq$) or ($\geq$) variety
  - we can move from one variety to the other (how?)
- Decision variables can be defined as non-negative (as in the previous example), non-positive, or free
- Decision variables in electric power systems:
  - Non-negative: production, demand, stored hydro energy in a hydroelectric unit
  - Free: flow of power on a transmission line, amount of power produced by a pumped hydro unit

Graphical solution of linear programming models

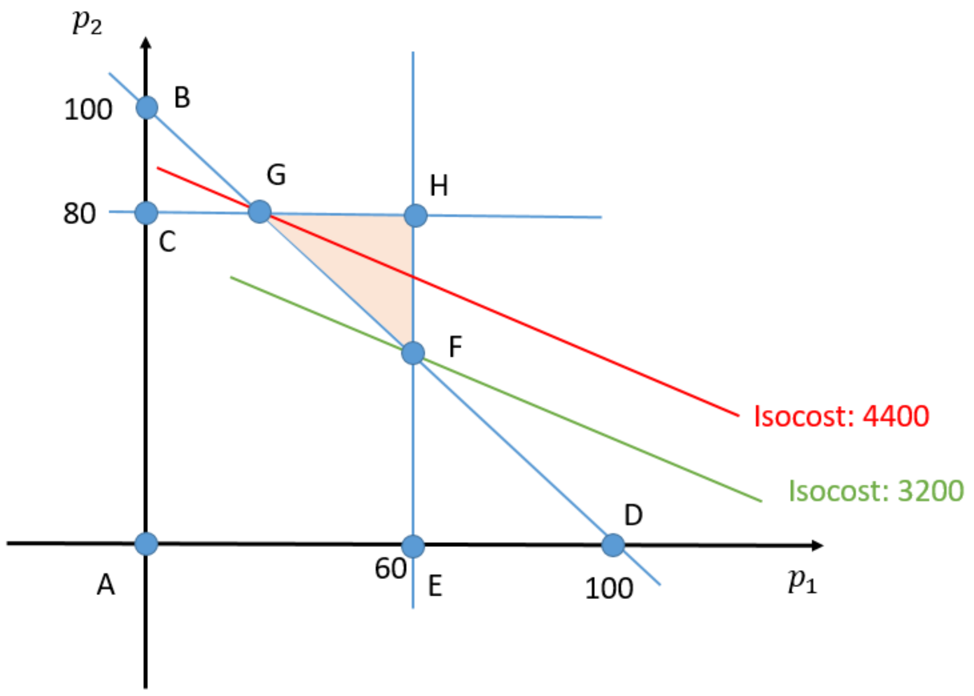
Graphical solution of economic dispatch problem: feasible set

- Feasible set: pink surface
- $p_1 \geq 0$ and $p_2 \geq 0$: the feasible set is in the non-negative orthant
- $p_1 + p_2 \geq 100$: half-space to the upper right of the line that crosses $(p_1, p_2) = (100, 0)$ and $(p_1, p_2) = (0{,}100)$
- $p_1 \leq 60$: half-scape to the left of the line $p_1 = 60$
- $p_1 \leq 80$: half-space under the line $p_2 = 80$



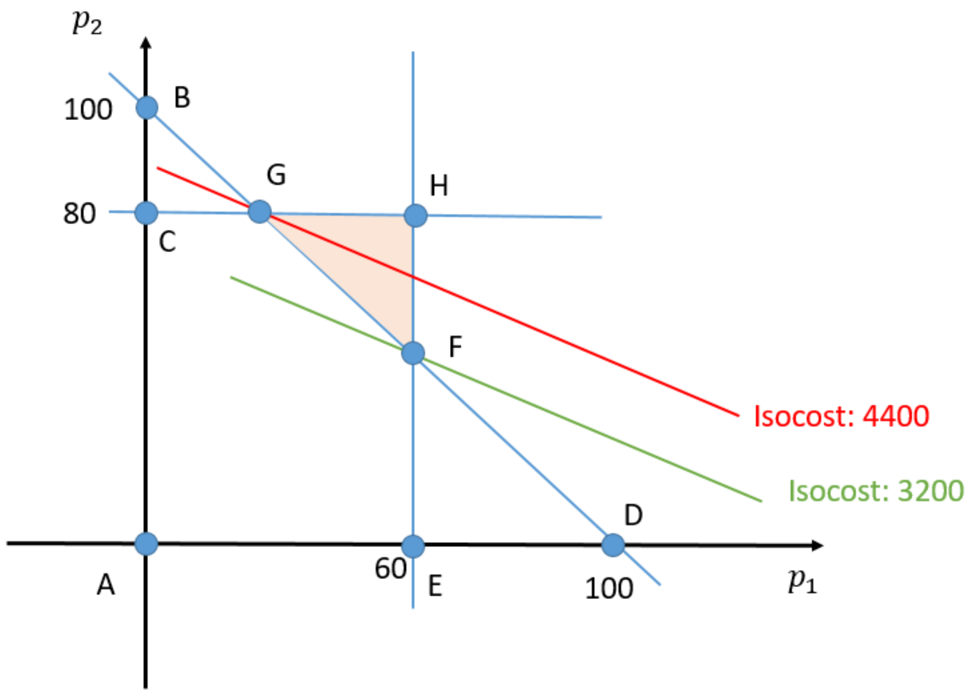
Graphical solution of economic dispatch problem: objective function

- Behavior of objective function: depicted by iso-cost lines of the objective function $z = 20 \cdot p_1 + 50 \cdot p_2$
- Iso-cost of \$4400: corresponds to the line $20 \cdot p_1 + 50 \cdot p_2 = 4400$ which is indicated in red
  - There are infinitely many points within the feasible set that achieve this cost (which ones?)
- Iso-cost of \$3200: parallel to the iso-cost of \$4400, indicated as the green line



Graphical solution of economic dispatch problem: optimal solution

- The green iso-cost is preferable to the red one
- It is as far down and to the left as the feasible set allows for
- The only point in the pink set that can attain this objective function value is F ⇒ optimal solution
- Coordinates of F: intersection of $p_1 = 60$ and $p_1 + p_2 = 100$, ⇒ $(p_1^*, p_2^*) = (60{,}40)$



Usefulness of graphical solution

- This two-dimensional analysis is quite limited, since we cannot visualize beyond three dimensions
- Nevertheless, it does clarify an important geometric intuition: the «corner points» of the feasible set have a special role in linear programs, because they are candidates for optimal solutions
- This geometric intuition is the basis of the celebrated simplex algorithm

Linear programs in standard form

Linear programs in standard form

A linear programming model in **standard form** with n decision variables and m equality constraints:

$$\min_x \sum_{i=1}^{n} c_i x_i$$

$$\text{s.t.} \sum_{i=1}^{n} A_{ij} x_i = b_j, j = 1, \ldots, m$$

$$x_i \geq 0, i = 1, \ldots, n$$

Vector and matrix notation

- The parameters c_i can be organized into a vector c, with transpose $c^T \in \mathbb{R}^{1\times n}$, that corresponds to the coefficients of the objective function of the linear program
- The parameters A_{ij} are coefficients of the problem constraints, and can be organized into a matrix $A \in \mathbb{R}^{m\times n}$ with m rows (one row per constraint) and n columns (one column per decision variable)
- The parameters b_j can be collected into a column vector $b \in \mathbb{R}^{m\times 1}$
- The problem can be expressed equivalently as:

$$\min_x c^T x$$

$$\text{s.t.} Ax = b$$

$$x \geq 0$$

Converting a linear program into standard form

- Any linear program can be expressed in standard form using the following transformations:
  - exchange maximization with minimization
  - introduce non-negative slack variables
- The transformation of linear programs into standard form allows us to better understand the relationship between extreme points of the feasible set and the underlying linear algebra computations that the simplex algorithm executes in order to solve the problem

Exchanging maximization with minimization

Exchanging maximization with minimization amounts to exchanging $\max_x c^T x$ with $\min_x - c^T x$

Introducing non-negative slack variables

- The use of non-negative **slack variables** allows us to transform any inequality to an equality by adding or subtracting non-negative quantities in such a way as to preserve an equivalence with the original inequality
- For example, the inequality $p_1 \leq 60$ can be expressed equivalently as $p_1 + s = 60$, where $s \geq 0$
  - This is equivalent to the original condition, because if we can find a non-negative variable $s \geq 0$ such that $p_1 + s = 60$, then it must be that $p_1 = 60 - s \leq 60$

Example A.4: economic dispatch problem in standard form

Returning to the economic dispatch problem, we can express it in standard form as follows:

$$\min_{p_1,p_2,s_1,s_2,s_3} 20 \cdot p_1 + 50 \cdot p_2 + 0 \cdot s_1 + 0 \cdot s_2 + 0 \cdot s_3$$

$$p_1 + p_2 - s_1 = 100$$

$$p_1 + s_2 = 60$$

$$p_2 + s_3 = 80$$

$$p_1, p_2, s_1, s_2, s_3 \geq 0$$

Extreme points and the simplex algorithm

Constraints are typically not more than decision variables

- Let us consider a linear program in standard form, and let us ignore the fact that variables are non-negative for the moment
- If there are more constraints (m equalities) than decision variables (n decision variables), then the linear system does not have a solution unless there are linearly dependent constraints
- Thus, we can focus on the case where there are more variables than constraints, $n \geq m$

Basis and basic solution

- If $m \leq n$, the set of feasible solutions has a dimension of up to $n - m$
- One way to attempt to compute feasible solutions is by setting $n - m$ of these variables equal to 0, and isolating the remaining $m \times m$ linear system
- Equivalent to splitting the original constraint matrix A into two parts, $A = [B\ N]$, where B is an $m \times m$ matrix and N is an $m \times (n - m)$ matrix
- If the remaining sub-matrix B, which is called the **basis**, is invertible, then the linear system has a unique solution
- This is called a **basic solution**

$$\Big[\overbrace{B}^{\substack{m\text{ columns}\\ \text{(basic variables)}}} \quad \underbrace{N}_{\substack{n-m\text{ columns}\\ \text{(non-basic variables)}}}\Big] \; \Big\} \; \substack{m\text{ lines}\\ \text{(constraints)}}$$

Basic and non-basic variables

- The $n - m$ variables that are set equal to zero at the outset are called **non-basic variables**
- The remaining m variables that are part of the linear system are called **basic variables**

Basic feasible solutions and extreme points

- If the unique solution of the $m \times m$ system is non-negative, then we have a **basic feasible solution** to the original problem
- An **extreme point** is a point that cannot be expressed as a convex combination of two other distinct points in the feasible set
- Geometrically, extreme points correspond to "corners" of the set of feasible solutions
- A basic feasible solution corresponds geometrically to an extreme point of the original set of feasible solutions

Example A.5: computing a basic feasible solution of the economic dispatch problem

- Consider the economic dispatch problem:

$$\min_{p_1,p_2,s_1,s_2,s_3} 20 \cdot p_1 + 50 \cdot p_2 + 0 \cdot s_1 + 0 \cdot s_2 + 0 \cdot s_3$$
$$p_1 + p_2 - s_1 = 100$$
$$p_1 + s_2 = 60$$
$$p_2 + s_3 = 80$$
$$p_1, p_2, s_1, s_2, s_3 \geq 0$$

- The constraint matrix of the problem can be expressed as:

$$A = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

- The columns of the matrix corresponds to the variables $(p_1, p_2, s_1, s_2, s_3)$ respectively

Example A.5: selecting non-basic variables

- The rank of the matrix is 3 (all constraints of the problem are linearly independent)
- Selecting (s_1, s_3) as the non-basic variables which are set equal to zero, the remaining linear system which only consists of the basic variables (p_1, p_2, s_2) can be expressed as:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ s_2 \end{bmatrix} = \begin{bmatrix} 100 \\ 60 \\ 80 \end{bmatrix}$$

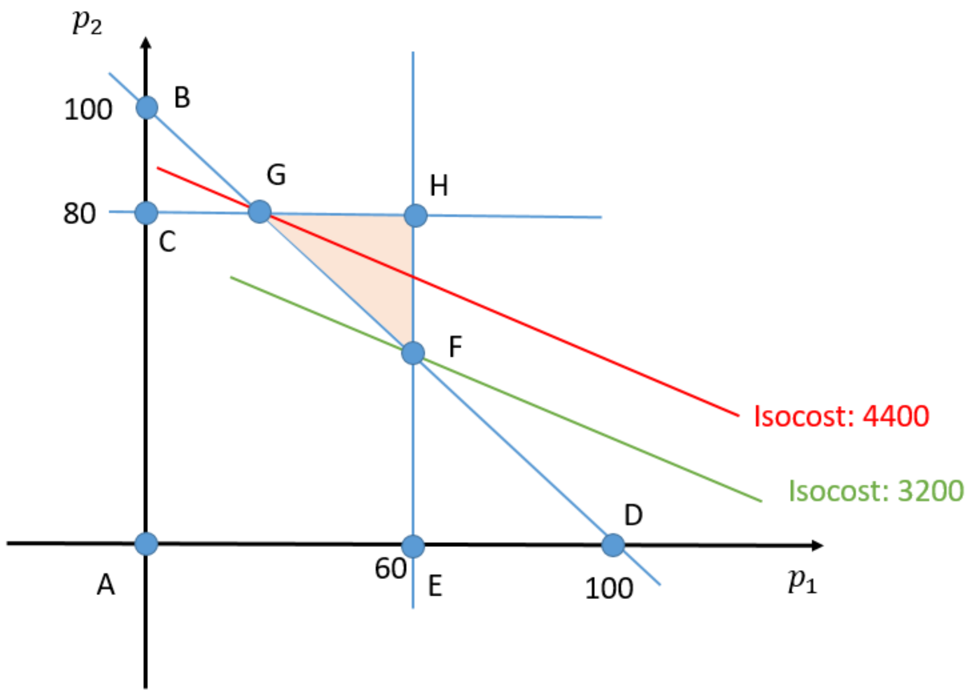
Example A.5: basic solution

- This system is obtained by isolating columns 1, 2 and 4 of the original matrix A
- The new 3×3 matrix is invertible because its rank is also 3
- Inverting the matrix, we compute the basic solution that corresponds to the basic variables (p_1, p_2, s_2)
- This corresponds to solving a system of 3 equations in 3 unknowns:

$$\begin{bmatrix} p_1 \\ p_2 \\ s_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 100 \\ 60 \\ 80 \end{bmatrix} = \begin{bmatrix} 20 \\ 80 \\ 40 \end{bmatrix}$$

Example A.5: basic optimal solution

- Non-negative basic variables ⇒ basic feasible solution
- Corresponds to the extreme point G
- $s_3 = 0 \Rightarrow$ the third constraint of the problem is tight at G ⇒ the point G is on the line $p_2 = 80$



One of the basic solutions is optimal

- One of the central ideas behind the simplex algorithm is that, if the original problem has an optimal solution (i.e. if it is neither infeasible nor unbounded), then one of these can be found among the extreme points of the feasible set
- Therefore, if we enumerate all extreme points of the feasible set and pick the best one, it is also an optimal solution
- There may also exist optimal solutions that are not extreme points, but at least one optimal solution is an extreme point
- In practice, enumerating all extreme points is computationally prohibitive: up to $\binom{n}{n-m}$ such points
- At least we can navigate between a finite number of points instead of infinitely many

Interlude: $\binom{n}{n-m}$

- The expression $\binom{n}{n-m} = \frac{n!}{n!(n-m)!}$ is the number of ways in which we can pick $n-m$ objects among n objects
- Or m objects among n objects: $\binom{n}{n-m} = \binom{n}{m}$
- In our case:
  - n is the number of decision variables
  - m is the number of basic variables (or $n-m$ is the number of non-basic variables)

Example A.6: basic solutions of the economic dispatch problem

- We return to the economic dispatch example and describe the set of basic solutions

$$\min_{p_1,p_2,s_1,s_2,s_3} 20 \cdot p_1 + 50 \cdot p_2 + 0 \cdot s_1 + 0 \cdot s_2 + 0 \cdot s_3$$

$$p_1 + p_2 - s_1 = 100$$

$$p_1 + s_2 = 60$$

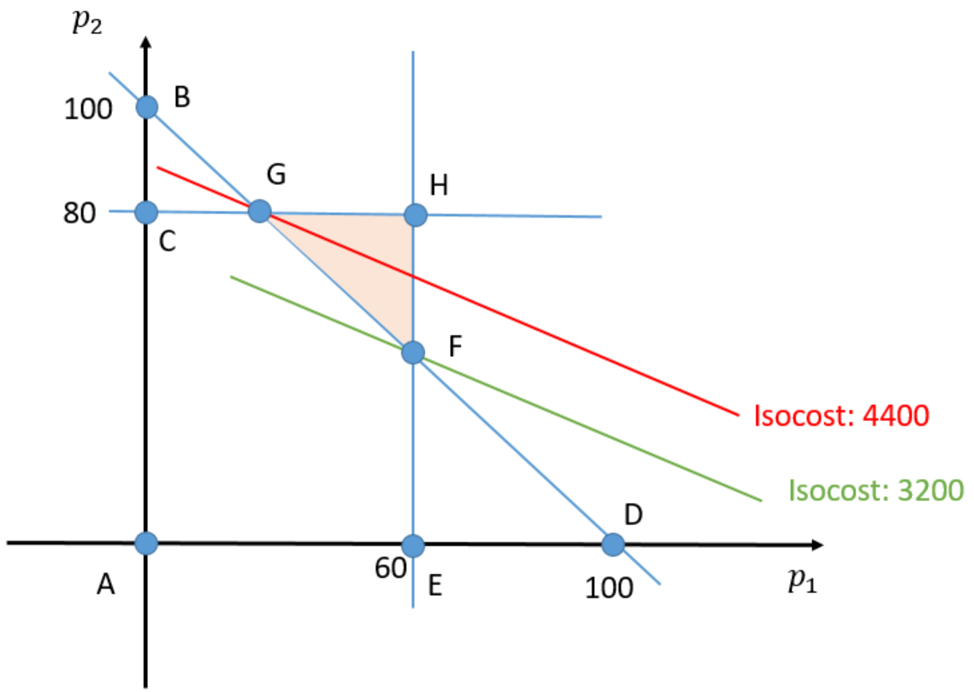
$$p_2 + s_3 = 80$$

$$p_1, p_2, s_1, s_2, s_3 \geq 0$$

- We consider two non-basic variables at a time, so that we are left over with a system of three variables in three constraints

Example A.6: matrix of basic variables

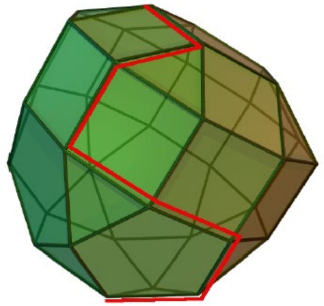
Non-basic variables	Basic variables	Basic solution	Extreme point	Feasible (≥ 0)?	Objective function
(p_1, p_2)	(s_1, s_2, s_3)	(-100, 60, 80)	A	No	-
(p_1, s_1)	(p_2, s_2, s_3)	(100, 60, -20)	B	No	-
(p_1, s_2)	(p_2, s_1, s_3)		-	-	-
(p_1, s_3)	(p_2, s_1, s_2)	(80, -20, 60)	C	No	-
(p_2, s_1)	(p_1, s_2, s_3)	(100, -40, 80)	D	No	-
(p_2, s_2)	(p_1, s_1, s_3)	(60, -40, 80)	E	No	-
(p_2, s_3)	(p_1, s_1, s_2)		-	-	-
(s_1, s_2)	(p_1, p_2, s_3)	(60, 40, 40)	F	Yes	3200
(s_1, s_3)	(p_1, p_2, s_2)	(20, 80, 40)	G	Yes	4400
(s_2, s_3)	(p_1, p_2, s_1)	(60, 80, 40)	H	Yes	5200



Example A.6: understanding the table

- There are $\binom{5}{2} = 10$ ways in which we can select non-basic variables
- There are 8 choices that lead to invertible sub-matrices, i.e. there are 8 basic solutions
- The choices that lead to non-invertible 3×3 sub-matrices are indicated with red font in the first column of the matrix, and in this case we cannot compute basic solution
  - Geometrically, they correspond to parallel lines that do not intersect
- The table presents the basic solutions, whether feasible or not, as well as the corresponding objective function value
- The basic feasible solutions correspond exactly to the corners of the feasible set and the optimal solution of the problem corresponds to one of these extreme points (point F, with a cost of $3200)

Idea of the simplex algorithm



- Instead of enumerating all the extreme points, the simplex algorithm navigates itself by applying on **pivot** per iteration
  - Algebraically, this corresponds to the exchange of a non-basic solution with a basic solution in the $m \times m$ linear system
  - Geometrically, this corresponds to jumping to a neighboring extreme point
- At every iteration (i.e. at every basic solution that is computed during the course of the algorithm), there is detailed theory regarding the numerical checks that are required for asserting if the current solution is optimal

Why the simplex algorithm works

- If the current basic solution is not optimal, there are processes for selecting a non-basic variable that should be removed from the basis, in a way that guarantees that the next iteration produces a solution that is **at least** as good as the present solution
- Given that the new solution is not worse than the previous one, and since we have a process for avoiding cyclic paths between solutions that attain the same objective function value,
  - Convergence: we have a guarantee that the process has a finite number of steps
  - Computational savings: we can avoid exploring a (potentially huge) number of points which have no hope of being better than the current iterate

Linear programming duality

Duality theory

- Duality theory is part of mathematical programming theory that concerns *all* classes of math programs (not only linear programming)
- It has important implications for algorithm development as well as the economic interpretation of mathematical programming models

Primal problem, dual problem and dual variables

- If the original problem, which is referred to as **primal problem**, is a minimization problem then the dual problem is a maximization, while if the primal is a minimization problem then the dual is a maximization problem
- The decision variables of the dual problem are determined by assigning a **dual variable** to each constraints of the primal problem
- To each decision variable of the primal problem corresponds a dual constraint

Duality mnemonic table

	Minimization	Maximization	
Constraints	$\geq b_i$	≥ 0	Variables
	$\leq b_i$	≤ 0	
	$= b_i$	Free	
Variables	≥ 0	$\leq c_j$	Constraints
	≤ 0	$\geq c_j$	
	Free	$= c_j$	

Dual of a problem in standard form

- Consider a primal problem in standard form:

$$\min_x c^T x$$

$$(\pi): Ax = b$$

$$x \geq 0$$

- Its dual is:

$$\max_\pi b^T \pi$$

$$\pi^T A \leq c^T$$

Example A.7: primal minimization probem

Consider the following linear program:

$$\min_x 3x_1 + 2x_2 + x_3$$

$$(\pi_1): x_1 + x_2 + x_3 \leq 3$$

$$(\pi_2): 2x_1 + 2x_2 + 2x_3 \geq 0$$

$$(\pi_3): x_3 = 5$$

$$x_1 \geq 0, x_2 \leq 0$$

Example A.7: dual

The dual problem is expressed as:

$$\max_{\pi} 3\pi_1 + 0\pi_2 + 5\pi_3$$

$$(x_1): \pi_1 + 2\pi_2 \leq 3$$

$$(x_2): \pi_1 + 2\pi_2 \geq 2$$

$$(x_3): \pi_1 + 2\pi_2 + \pi_3 = 1$$

$$\pi_1 \leq 0, \pi_2 \geq 0$$

Example A.8: economic dispatch

- Consider the economic dispatch problem with a slight rearrangement of the objective function and first primal constraint:

$$\max_{p_1,p_2} -20 \cdot p_1 - 50 \cdot p_2$$

$$(\lambda): -p_1 - p_2 \leq -100$$

$$(\mu_1): p_1 \leq 60$$

$$(\mu_2): p_2 \leq 80$$

$$p_1, p_2 \geq 0$$

- Optimal objective function value: -3200

Example A.8: dual of the economic dispatch problem

- The dual problem can be expressed as:

$$\min_{\lambda,\mu} -100\lambda + 60\mu_1 + 80\mu_2$$

$$(p_1): -\lambda + \mu_1 \geq -20$$

$$(p_2): -\lambda + \mu_2 \geq -50$$

$$\lambda \geq 0, \mu_1 \geq 0, \mu_2 \geq 0$$

- The optimal solution of the dual problem is $\lambda^* = 50,\ \mu_1^* = 30,\ \mu_2^* = 0$
- The optimal objective function value of the dual problem is -3200 (equal to the optimal objective function value of the primal problem)

Relaxation

- A **relaxation** of a mathematical program is a variation of the original mathematical program in which the feasible solution is expanded
  - For instance, by ignoring some constraint
- Thus, a relaxation of a minimization problem yields a solution that results in an objective function value that is less then or equal to the objective function of the original minimization problem
- Similarly, a relaxation of a maximization problem yields a solution that results in an objective function value that is greater than or equal to the objective function value of the original maximization problem

Weak duality

- Dual problems result in a bound for corresponding primal problems, because they come from (non-trivial) *relaxations* of the original primal problems
- This is referred to as **weak duality**, and holds for all mathematical programs

Strong duality

- A stronger result in mathematical programming states that, for certain classes of math programs, the objective function value of the dual problem becomes *equal* to the objective function value of the primal problem
- This is referred to as **strong duality**
- Strong duality holds in the case of linear programs
- It is often the case that math programs for which strong duality holds are computationally manageable

Strong duality in linear programming

- Strong duality in linear programming can be expressed in a more nuanced way, in order to capture cases where linear programs are infeasible or unbounded:
  - If a primal problem has an optimal solution, then its dual has an optimal solution with an objective function value that is equal to that of the primal problem
  - If a primal problem is unbounded, then its dual is infeasible
  - If a primal problem is infeasible, then its dual may be infeasible or unbounded

Usefulness of duality

- Duality is used in optimization algorithms for solving large-scale problems
- The general principle is that one solves the dual problem, which is typically easier to solve than the original primal problem
- This leads to a bound of the original problem in the case of weak duality (from which we can sometimes extract high-quality primal solutions), or to the optimal solution of the original problem in the case of strong duality

Dual of the dual

The dual of a dual problem corresponds to the primal problem

Sensitivity

Sensitivity and dual variables

- Dual variables hold significant information regarding the behavior of the objective function at the neighborhood of the optimal solution
- Specifically, a dual variable quantifies the sensitivity of the objective function of the original problem to a change in the right-hand side of the corresponding constraint

Example A.9: economic dispatch

- Consider the economic dispatch problem, where we increase the right-hand side of the first constraint ($-p_1 - p_2 \leq -100$) by one, thus converting the right-hand side of the constraints to -99 instead of -100:

$$\max_{p_1,p_2} -20 \cdot p_1 - 50 \cdot p_2$$

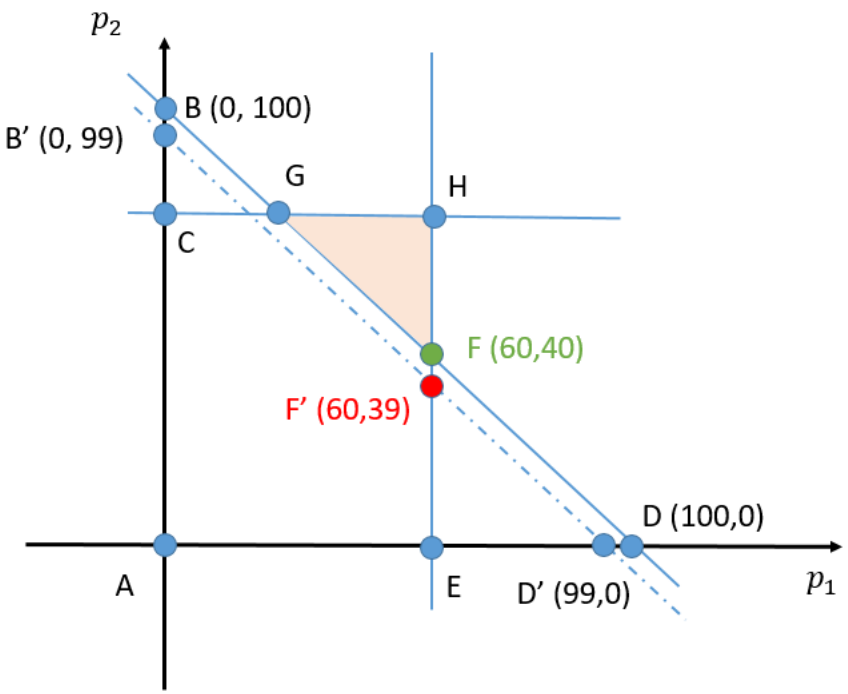
$$(\lambda): -p_1 - p_2 \leq -99$$

$$(\mu_1): p_1 \leq 60$$

$$(\mu_2): p_2 \leq 80$$

$$p_1, p_2 \geq 0$$

Example A.9: graphical solution



Example A.9: graphical solution of the new problem

- Point F corresponds to the original optimal solution and point F’ corresponds to the new optimal solution
- The feasible set has increased, since the constraint $p_1 + p_2 \geq 100$ has shifted to the lower left, and is now expressed as $p_1 + p_2 \geq 99$
- This constraint is indicated by the line that crosses points B' (0, 99) and D' (99,0)
- The expansion of the feasible set leads to a new and cheaper iso-cost that crosses F’ at (60, 39)
- The cost of this solution is 3150, this the optimal objective function value of the problem expressed as a maximization is -3150
- The change in the optimal value of the objective function is therefore equal to -3150-(-3200)=50
- We observe that this is exactly the value of the optimal dual value λ^*, which is the dual variable of the constraint that we have relaxed

KKT conditions for linear programs

Karush-Kuhn-Tucker (KKT) conditions

- The **Karush-Kuhn-Tucker (KKT) conditions** are a set of mathematical conditions that characterize the optimal solution of a primal problem and its dual for certain classes of mathematical programs, including linear programs
- The KKT conditions are a set of inequalities and **complementarity conditions** that implicate both primal as well as dual variables, and which serve as certificates for the optimality of the primal and dual poroblems
- They are *necessary* and *sufficient* for certain classes of problems (including linear programs):
  - If a primal-dual vector (x, π) is a candidate for optimality, this can be checked by testing whether it satisfies the KKT conditions (sufficient)
  - Any primal optimal solution x and dual optimal solution π must satisfy these conditions (necessary)

Complementarity operator

- The complementarity operator is indicated as $\perp$
- The expression $a \perp b$ implies that $a \cdot b = 0$
- We also use the condensed notation $0 \leq a \perp b \geq 0$ to indicate that the following three conditions hold simultaneously:

$$a \geq 0$$
$$b \geq 0$$
$$a \cdot b = 0$$

- Thus, $0 \leq a \perp b \geq 0$ implies that either $a = 0$, or , $b = 0$, or both, but both quantities cannot be positive simultaneously, therefore if one is positive then the other is zero
- The notation is generalized to vectors

KKT conditions for maximization linear programs

- Consider the following linear program:

$$\max_{x,y} c_x^T x + c_y^T y$$
$$\text{s.t. } (\lambda): Ax + By \leq b$$
$$(\mu): Cx + Dy = d$$
$$x \geq 0$$

- The KKT conditions of the problem have the following form:

$$Cx + Dy - d = 0$$
$$0 \leq \lambda \perp Ax + By - b \leq 0$$
$$0 \leq x \perp \lambda^T A + \mu^T C - c_x^T \geq 0$$
$$\lambda^T B + \mu^T D - c_y^T = 0$$

and are necessary and sufficient for an optimal solution

KKT conditions for minimization linear programs

- Consider the following linear program:

$$\min_{x,y} c_x^T x + c_y^T y$$
$$\text{s.t. } (\lambda): Ax + By \leq b$$
$$(\mu): Cx + Dy = d$$
$$x \geq 0$$

- The KKT conditions of the problem have the following form:

$$Cx + Dy - d = 0$$
$$0 \leq \lambda \perp Ax + By - b \leq 0$$
$$0 \leq x \perp \lambda^T A + \mu^T C + c_x^T \geq 0$$
$$\lambda^T B + \mu^T D + c_y^T = 0$$

and are necessary and sufficient for an optimal solution

Example A.10: economic dispatch

Consider the economic dispatch problem:

$$\max_{p_1,p_2} -20 \cdot p_1 - 50 \cdot p_2$$

$$(\lambda): -p_1 - p_2 \leq -100$$

$$(\mu_1): p_1 \leq 60$$

$$(\mu_2): p_2 \leq 80$$

$$p_1, p_2 \geq 0$$

Example A.10: economic dispatch

The KKT conditions of the problem are summarized as follows:

$$0 \leq \lambda \perp p_1 + p_2 - 100 \geq 0$$

$$0 \leq \mu_1 \perp 60 - p_1 \geq 0$$

$$0 \leq \mu_2 \perp 80 - p_2 \geq 0$$

$$0 \leq p_1 \perp 20 - \lambda + \mu_1 \geq 0$$

$$0 \leq p_2 \perp 50 - \lambda + \mu_2 \geq 0$$

Example A.10: confirming the KKT conditions

- We claimed earlier that $\lambda^* = 50, \mu_1^* = 30, \mu_2^* = 0$ is an optimal solution to the dual problem, without however proving it
- This can be confirmed by checking that the candidate dual solution satisfies the KKT conditions of the economic dispatch problem when combined with the primal solution $p_1^* = 60, p_2^* = 40$

Usefulness of the KKT conditions

- The KKT conditions are used extensively in the book, because they allow us to extract quantitative conclusions regarding perfect competition models
- Economic models of perfect competition are the starting point for modeling electricity markets
- The KKT conditions thus characterize the behavior of market prices in conditions of perfect competition, and explain the behavior of a large range of models:
  - pricing energy in economic dispatch models
  - pricing network access in optimal power flow problems
  - pricing reserves in energy and reserve co-optimization problems
  - pricing capacity in long-term investment models
  - pricing energy in ramp-constrained models
  - effect of storage on market prices
  - effect of substitution on energy prices
  - ...

Example A.11: economic interpretation of dual variables

- Suppose that, instead of receiving an instruction about the amount of energy production, generator 1 responds to a market price, which we denote as λ
- Given an exogenous price λ, the problem of maximizing generator profit can be expressed as follows:

$$\max_{p_1 \geq 0} \lambda \cdot p_1 - 20 \cdot p_1$$

$$(\mu_1): p_1 \leq 60$$

- The choice of notation follows that of the economic dispatch problem

Example A.11: maximizing generator profit

- The KKT conditions of this profit maximization problem are expressed as follows:

$$0 \leq p_1 \perp 20 - \lambda + \mu_1 \geq 0$$

$$0 \leq \mu_1 \perp 60 - p_1 \geq 0$$

- We observe that these conditions are *identical* to the second and fourth KKT condition of the original economic dispatch problem
- This is an important observation: it implies that the optimal solution of the economic dispatch problem *contains* the profit maximization goal of generator 1
- In other words, this means that the primal-dual solution $(\lambda^*, \mu_1^*, p_1^*)$ that is produced by the economic dispatch model also solves the profit maximization problem of generator 1, since the dual variable λ assumes the role of *market price*
- Similarly, the profit maximization problem of generator 2 is contained in the second and fifth KKT condition of the economic dispatch problem

Example A.11: market clearing

- This leaves us with interpreting the first KKT condition of the economic dispatch problem: $0 \leq \lambda \perp p_1 + p_2 - 100 \geq 0$
- This condition is interpreted as a *market clearing condition*:
  - given a non-zero market price ($\lambda > 0$), the production of units should equal the market demand ($p_1 + p_2 = 100$)
  - unless there is excess production in the market ($p_1 + p_2 > 100$), in which case the energy price is zero ($\lambda = 0$) because there is over-supply

Example A.11: overall interpretation of KKT conditions

- Thus, the KKT conditions of the centralized economic dispatch problem can be considered to be equivalent to a set of conditions that contain the following information:
  - The primal-dual solution must be such that the profit of generator 1 is maximized
  - The primal-dual solution must be such that the profit of generator 2 is maximized
  - The market clears

Example A.12: pricing with ramp constraints

- Consider an economic dispatch problem in two time periods in a system that consists of two generators
- Load:
  - Period 1: 100 MWh
  - Period 2: 200 MWh

Generator	Marginal cost ($/MWh)	Ramp rate (MW)	Capacity (MW)
1	20	60	$+\infty$
2	50	$+\infty$	$+\infty$

Example A.12: linear programming model

The linear program that describes the optimal dispatch can be expressed as follows:

$$\min_{p \geq 0} 20 \cdot (p_{11} + p_{12}) + 50 \cdot (p_{21} + p_{22})$$

$$(\lambda_1): 100 - p_{11} - p_{21} = 0$$

$$(\lambda_2): 200 - p_{12} - p_{22} = 0$$

$$(\delta^+): p_{12} - p_{11} \leq 60$$

$$(\delta^-): p_{11} - p_{12} \leq 60$$

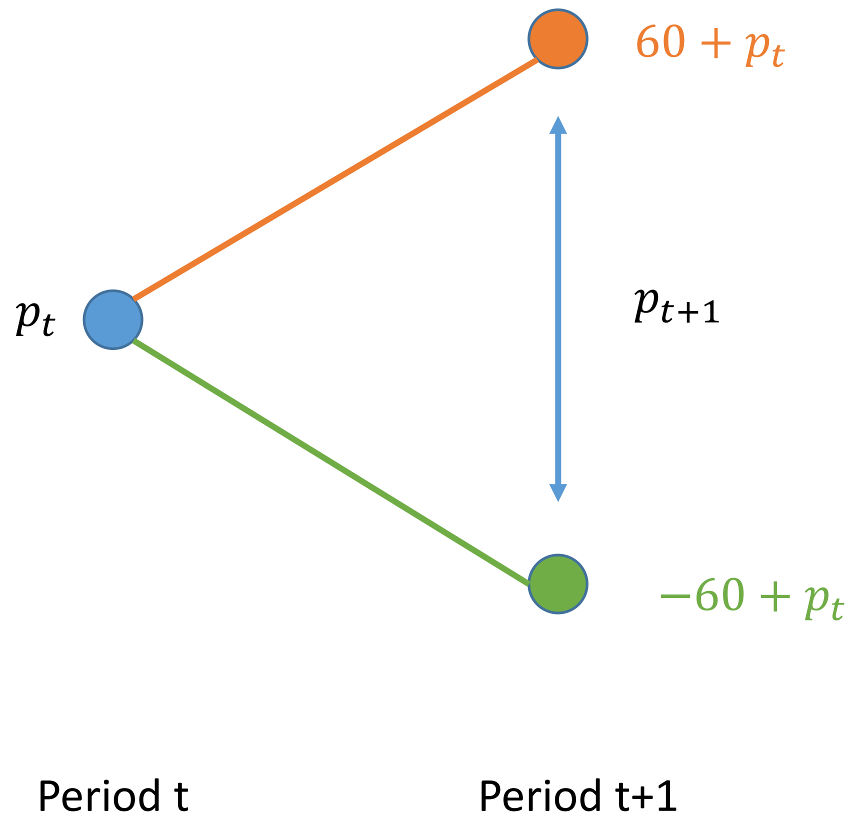
Example A.12: ramp constraints

- Ramp constraints limit how much the output of a generator can change from one time period to the next:

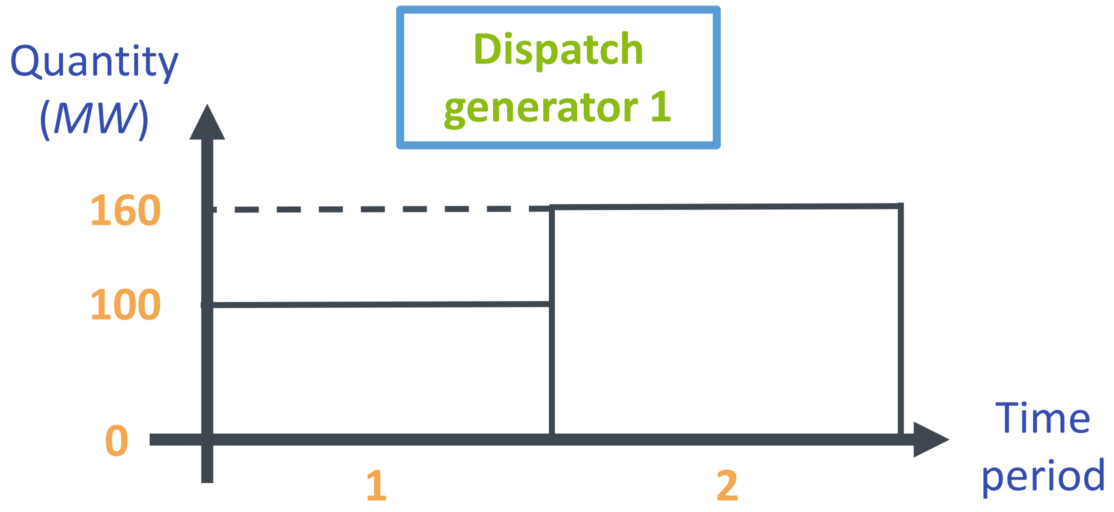
$$p_{t+1} - p_t \leq 60$$
$$p_t - p_{t+1} \leq 60$$

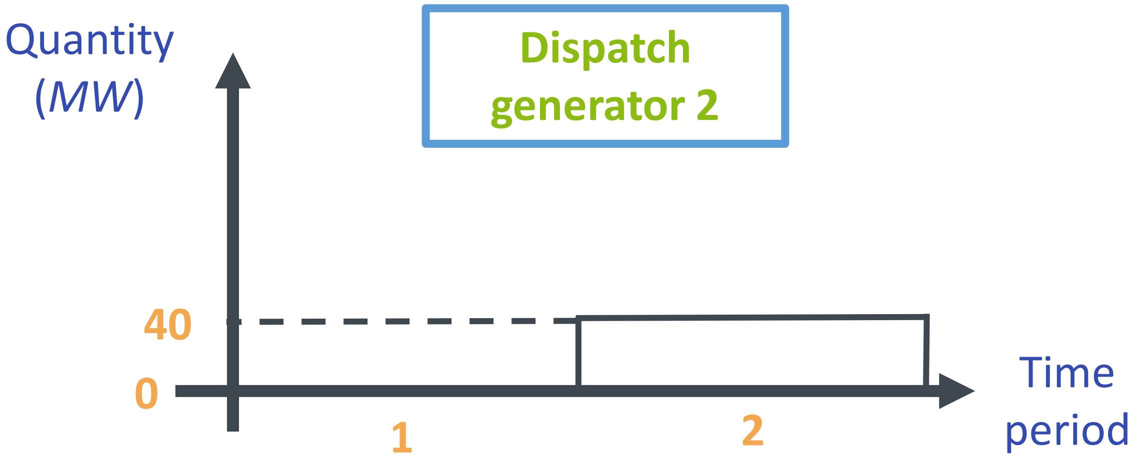
- Equivalent to:

$$p_{t+1} \leq 60 + p_t$$
$$-60 + p_t \leq p_{t+1}$$



Example A.12: optimal solution





Example A.12: KKT conditions

The KKT conditions of the problem are described as follows:

$$100 - p_{11} - p_{21} = 0$$
$$200 - p_{12} - p_{22} = 0$$

Market clearing

$$0 \leq 60 - p_{12} + p_{11} \perp \delta^+ \geq 0$$
$$0 \leq 60 - p_{11} + p_{12} \perp \delta^- \geq 0$$
$$0 \leq 20 - \lambda_1 - \delta^+ + \delta^- \perp p_{11} \geq 0$$
$$0 \leq 20 - \lambda_2 + \delta^+ - \delta^- \perp p_{12} \geq 0$$

Profit maximization generator 1

$$0 \leq 50 - \lambda_1 \perp p_{21} \geq 0$$
$$0 \leq 50 - \lambda_2 \perp p_{22} \geq 0$$

Profit maximization generator 2

Example A.12: profit maximization of generator 1

- The green KKT conditions are equivalent to the following profit maximization problem of generator 1:

$$\max_{p \geq 0} (\lambda_1 - 20) \cdot p_{11} + (\lambda_2 - 20) \cdot p_{12}$$

$$(\delta^+): p_{12} - p_{11} \leq 60$$

$$(\delta^-): p_{11} - p_{12} \leq 60$$

Example A.12: profit maximization of generator 2

- The orange KKT conditions are equivalent to the following profit maximization problem of generator 2:

$$\max_{p \geq 0} (\lambda_1 - 50) \cdot p_{21} + (\lambda_2 - 50) \cdot p_{22}$$

Example A.12: understanding the price of period 2

- The profit maximization of each unit is contained in the KKT conditions of the centralized economic dispatch problem
- This observation can be used in order to understand the market prices that result from the model
- Specifically, since generator 2 is asked to produce a positive quantity in period 2, the only market price that can urge the unit to produce a non-zero but finite amount of energy in period 2 is its marginal cost, because
  - a price that is greater than its marginal cost would urge the unit to produce at its maximum (i.e. infinitely much)
  - while a price signal that is lower than its marginal cost would urge the unit to produce zero
- On the other hand, we cannot conclude what the price is for period 1 from the optimal dispatch of generator 2: since the generator is not producing, the price λ_1 must be less than or equal to its marginal cost, but understanding the precise value of the price requires the analysis of the profit maximization of generator 1

Example A.12: understanding the price in period 1

- If the average value of the price over the two periods is greater than 20 $/MWh, then generator 1 has an interest to produce an arbitrarily large amount of energy in period 1, and this same quantity plus 60 MWh (the ramp constraint) in period 2
- On the other hand, if the average value in the two periods is lower than 20 $/MWh, then the unit has an interest in producing 0 MWh during both periods
- Since none of these two extremes is the optimal dispatch, the average price in both periods must equal 20 $/MWh, so that we can urge the generator to produce 100 MWh in period 1 and 160 MWh in period 2
- Which implies that the price in period 1, λ_1, must equal -10 $/MWh
- A negative price may appear as being exotic, but it occurs in electricity markets, and implies that consumers *are paid*, instead of paying, to consume energy

Example A.12: inferring the price of period 1 from the KKT conditions

- Since $p_{11} > 0$, the price of period 1 is
$$\lambda_1 = 20 - \delta^+ + \delta^-$$
- Since $p_{12} > 0$ then
$$\lambda_2 = 20 + \delta^+ - \delta^-$$
- We already know that $\lambda_2 = 50$ \$/MWh
- Which implies that $\delta^+ - \delta^- = 30$ \$/MWh
- Which implies that $\lambda_1 = -10$ \$/MWh

Example A.12: (incorrect) pricing in practice

- It is worth juxtaposing the price derived previously to a heuristic that is used in practice by certain system operators: pricing at the marginal cost of the cheapest unit that is producing a non-zero quantity
- Applying this heuristic pricing method to the optimal solution of the economic dispatch that is calculated previously implies a price of 20 \$/MWh in both periods (since generator 1 produces a non-zero quantity in both periods, and is the cheapest unit in the market)
- However, this heuristic method does *not* maximize generator profits, and is therefore not aligned with private incentives

Representing piecewise linear functions

Representing piecewise linear functions

Hyperplane in $\mathbb{R}^n$: a set of points that is described as

$$\{x \in \mathbb{R}^n | a^T x = b\}$$

where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$

- Generalizes a «line» into higher dimensions

- A piecewise linear **convex** function corresponds to the **upper** envelope of the hyperplanes that define it
- A piecewise linear **concave** function corresponds to the **lower** envelope of the hyperplanes that define it

Example A.15: minimizing piecewise linear convex functions

- Consider the following non-linear problem
$$\min_x 2 \cdot x + |x| + 2.5 \cdot |x - 2|$$
- Optimal solution: $x = 0$
- Geometric «solution»: moving from left to right, we stop at the point where the slope switches from negative to positive

Example A.15: equivalent representation as a linear program

Equivalent representation as a linear program:

$$\min_{x,\theta_1,\theta_2} 2 \cdot x + \theta_1 + 2.5 \cdot \theta_2$$

$$\theta_1 \geq x$$

$$\theta_1 \geq -x$$

$$\theta_2 \geq x - 2$$

$$\theta_2 \geq 2 - x$$

Piecewise linear functions and dynamic programming

- Dynamic programming **value functions** in multistage stochastic linear programs are piecewise linear
- They emerge in the context of hydrothermal planning
- They map the level of water to the expectation of the value function

Sensitivity analysis

Condition for an optimal basis

<u>Proposition</u>: A basic solution is optimal if

- $B^{-1}b \geq 0$ and
- $c_N^T - c_B^T B^{-1} N \geq 0$

- The solution is feasible because $x_B = B^{-1}b \geq 0$ and $x_N = 0$
- We express the basic variables as a function of non-basic variables in the objective function:

$$[B \quad N]\begin{bmatrix} x_B \\ x_N \end{bmatrix} = b \Leftrightarrow Bx_B + Nx_N = b \Leftrightarrow x_B = B^{-1}(b - Nx_N)$$

- Substituting into the objective function:

$$c^T x = c_B^T B^{-1} b + \left(c_N^T - c_B^T B^{-1} N\right) x_N$$

Condition for an optimal basis

- The first term is a constant
- If we move from the present solution to any other feasible solution, then the second term cannot improve the objective function
- The vector $c_N^T - c_B^T B^{-1} N \in \mathbb{R}^{n-m}$ is referred to as the **reduced cost** of the basis
- Non-negative reduced cost is a sufficient condition for a feqsaible basis to be optimal
- And under certain conditions (non-degenerate solution) it is also sufficient (theorem 3.1 [2])

Example A.16: Feasible and optimal bases for the diet problem

- We are seeking a least-cost diet which consists of three dishes and includes at least b_1 and b_2 units of two nutrients
- The content of dishes in nutrients is presented in the table
- Cost of dishes:
  - Dish 1: 1 €
  - Dish 2: 2 €
  - Dish 3: 1 €

	Dish 1	Dish 2	Dish 3
Nutrient 1	0.5	4	1
Nutrient 2	2	1	2

Example A.16: formulation as a linear program

The problem is expressed parametrically with respect to nutrient content as follows:

$$\min_x x_1 + 2 \cdot x_2 + x_3$$
$$0.5 \cdot x_1 + 4 \cdot x_2 + x_3 = b_1$$
$$2 \cdot x_1 + x_2 + 2 \cdot x_3 = b_2$$
$$x_1, x_2, x_3 \geq 0$$

Three possible bases:

$$B_1 = \begin{bmatrix} 0.5 & 4 \\ 2 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0.5 & 1 \\ 2 & 2 \end{bmatrix}, B_3 = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

Example A.16: basic solution as a function of nutritional requirements

- Reduced costs:
  - Basis B_1: -0.2
  - Basis B_2: 1.5
  - Basis B_3: 0.2143
- Basic solutions as functions of b:

$$x_{B_1} = \begin{bmatrix} -0.1333 \cdot b_1 + 0.5333 \cdot b_2 \\ 0.2667 \cdot b_1 - 0.0667 \cdot b_2 \end{bmatrix}$$

$$x_{B_2} = \begin{bmatrix} -2 \cdot b_1 + b_2 \\ 2 \cdot b_1 - 0.5 \cdot b_2 \end{bmatrix}$$

$$x_{B_3} = \begin{bmatrix} 0.2857 \cdot b_1 - 0.1429 \cdot b_2 \\ -0.1429 \cdot b_1 + 0.5714 \cdot b_2 \end{bmatrix}$$

Example A.16: cost as a function of b

- Define $R_i = \{(b_1, b_2): x_{B_i} \geq 0\}$ as the set of b for which basis i is feasible

$$R_1 = \{0.25 \cdot b_1 \leq b_2 \leq 4 \cdot b_1\}$$

$$R_2 = \{2 \cdot b_1 \leq b_2 \leq 4 \cdot b_2\}$$

$$R_3 = \{0.25 \cdot b_1 \leq b_2 \leq 2 \cdot b_1\}$$

- Cost of each basic solution as a function of the parameters (b_1, b_2)

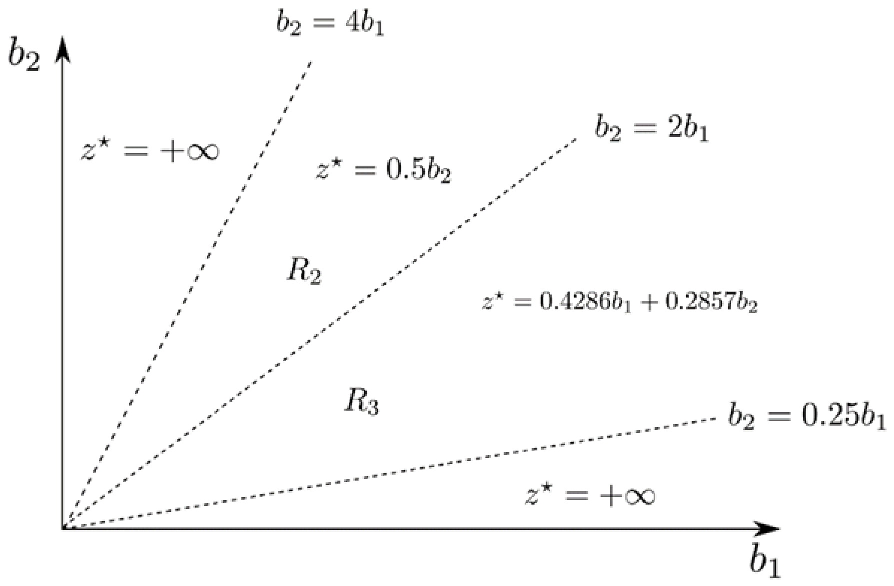
$$c_{B_1}^T x_{B_1} = 0.4 \cdot b_1 + 0.4 \cdot b_2$$

$$c_{B_2}^T x_{B_2} = 0.5 \cdot b_2$$

$$c_{B_3}^T x_{B_3} = 0.4286 \cdot b_1 + 0.2857 \cdot b_2$$

Example A.16: the function $z^*(b)$

- We can check that:
  - x_{B_2} is optimal in R_2
  - x_{B_3} is optimal in R_3
- The optimal solution as a function of b is piecewise linear
- The region R_2 corresponds to a mix of dishes 1 and 3
- The region R_3 corresponds to a mix of plates 2 and 3



$z^*(b)$ from a dual point of view

- The dual of the primal program in standard form is

$$(D): \max_\pi \pi^T b$$
$$\text{s.t.}\, \pi^T A \leq c^T$$

- A basic solution of a polyhedron $P \subseteq \mathbb{R}^n$ (not in standard form) is a vector x such that:
  - The equality constraints are active, and
  - From the equality constraints that are active at x, n of them are linearly independent
- Each base B of the constraint matrix A corresponds to a basic solution of the feasible set of the dual problem based on the relation $\pi^T = c_B^T B^{-1}$
- And for each basic solution of the dual, there is a basis in the primal matrix such that the above relationship holds
- Thus, the dual problem is expressed equivalently as

$$z^*(b) = \max_{i=1,\ldots,r} \pi_i^T b$$

where r: number of basic feasible solutions of (D)

- As in the primal analysis, this suggests that $z^*(b)$ is a piecewise linear function of b

Example A.17: feasible set of the dual of the diet problem

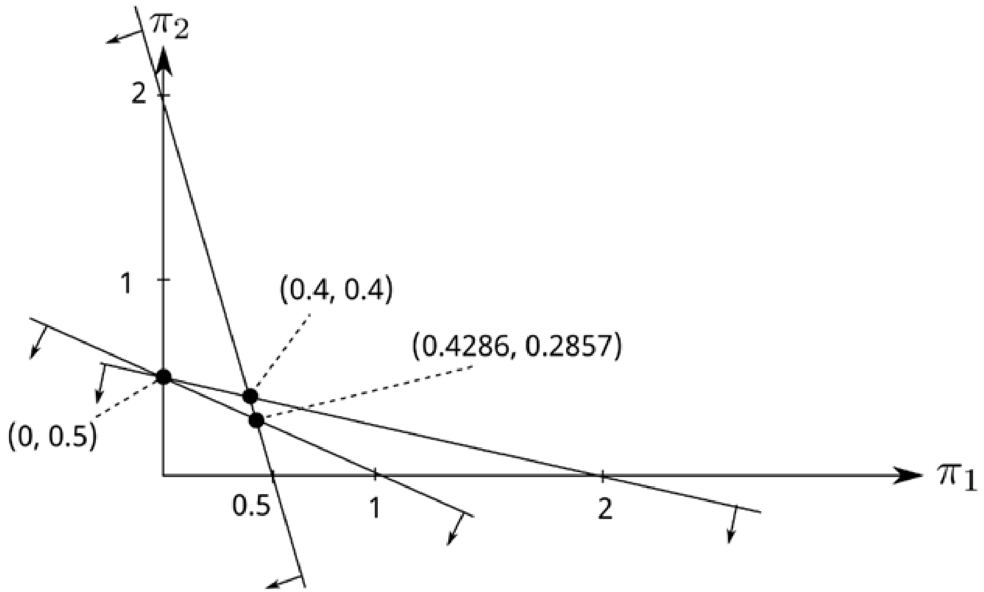
The dual of the diet problem is expressed as follows:

$$\max_{\pi} b_1 \cdot \pi_1 + b_2 \cdot \pi_2$$

$$\text{s.t. } 0.5 \cdot \pi_1 + 2 \cdot \pi_2 \leq 1$$

$$4 \cdot \pi_1 + \pi_2 \leq 2$$

$$\pi_1 + 2 \cdot \pi_2 \leq 1$$



Βιβλιογραφία

[1] A. Papavasiliou, Optimization Models in Electricity Markets, Cambridge University Press

https://www.cambridge.org/highereducation/books/optimization-models-in-electricity-markets/0D2D36891FB5EB6AAC3A4EFC78A8F1D3#overview

[2] Bertsimas, D. & Tsitsiklis, J. N., 1997. Introduction to Linear Optimization. Cambridge, MA: Athens Scientific