Chalmers University of Technology   MVE165
University of Gothenburg   MMG630
Mathematical Sciences   Applied Optimization
Optimization   Assignment information
Elin Svensson
Ann-Brith Strömberg   May 3, 2010

Assignment 3b: Investment planning

Given below is the description of the assignment Investment Planning. An implementation of the model in AMPL is found on the course homepage:
`http://www.math.chalmers.se/Math/Grundutb/CTH/mve165/0910`
The file `co2.run` should be edited in order to solve the different instances of the model, as described in the exercises below.

To pass the assignment you should (in groups of two persons) **(i)** write a **detailed report** that gives satisfactory answers and explanations to the exercise questions. The report should be e-mailed to `anstr@chalmers.se`
     at the latest on Monday 17 of May 2010 at 17.00.

Your shall also **(ii)** present your project orally at a seminar
     on Thursday 20 or Friday 21 of May 2010.

You are strongly adviced to **start working on the project exercises in good time before the deadline**! The solution time for the solver CPLEX (in the stochastic case) is usually between 30 seconds and three minutes. For the construction of Pareto graphs, several such solutions must be computed, which quickly leads to total solution times of 15–20 minutes.

We also ask you to estimate the number of hours you spent on this assignment and note this in your report or presentation.

Problem background

This project concerns the planning of energy-efficiency investments in a pulp mill. These investments consist of measures to reduce the steam consumption and measures that make use of the acquired steam surplus to enable exports of electricity, district heating, and/or the wood-fuel called lignin. Furthermore, the pulp mill is assumed to face a production increase, making certain capacity-increasing investments necessary.

The main objective is to maximize the expected net present value of the investments, that is, to find the economically optimal investment plan. There are also some special difficulties to consider:

- *Uncertainties* in future energy prices and associated CO_2 emissions should be considered.

- Not only economy, but also the resulting reductions of CO_2 emissions should be optimized. This yields a *multiobjective* programming model.

- Many of the decision variables are *binary*, representing whether an investment is made or not.

Multistage stochastic programming

Investment decisions are assumed to be made "here-and-now", before uncertainties about the energy market are resolved and before any changes in, for example, energy prices occur. A point in time when investment decisions can be made is followed by a period of five years, when no new investments can be made. The cash flow of the second stage (which is the period when no investments can be made) is a function of the previous investment decisions, the energy prices, and the operative decisions. If, for example, investments are made in a steam turbine in the first stage, the income in the second stage depends on the size of the turbine, the electricity price, and the amount of steam that is passed through the turbine.

After each investment period, new investments can be made, which are then followed by new periods with realizations of uncertain parameters and changed cash flows. A model of this kind, with two types of decisions where one is a reaction to the other as well as to the realization of the uncertain parameters, is termed a *multistage recourse model*. The uncertain parameters are modelled using *scenarios*.

CO_2 emissions in a systems perspective

Improving energy efficiency leads to global reductions of CO_2 emissions. For example, by reducing the use of fossil fuels, the emissions are directly reduced on-site. The reduction of biomass use will also lead to CO_2 emissions reductions, but in this case off-site. The reason is that biomass is assumed to be a limited resource, and hence by using less of it at one plant more can be used to substitute fossil fuels elsewhere, thereby reducing the overall emissions. Also reduced imports or increased exports of electricity affect the net CO_2 emissions.

This line of reasoning makes the assumptions about the surrounding energy market very important for an application like this. Assume, for example, that the mill can increase its electricity production. The produced electricity enables decreased electricity production somewhere else in the grid. Today this reduction will take place in a coal-fired condensing power plant. The reduction in emissions will be quite high. In the future, the reduction might be in a natural gas combined cycle plant, or in a coal-fired condensing power plant with a carbon capture unit, in which case the emissions reductions will be substantially lower. In this model, these assumptions are handled in the scenario model. The scenario model used here is built from three different blocks of energy market data, each based on different assumptions about the energy market prices and emissions.

The mathematical model

Only some important parts of the mathematical model are described here. A complete description of the model, including all constraints and variables, can be found in the article "A Scenario-Based Stochastic Programming Model for the Optimization of Process Integration Opportunities in a Pulp Mill", which can be found on the course homepage.

There have also been some changes to the model used in this project compared to the one used in the article. This concerns the scenario model and the multiple objectives, and hence those parts of the model will be described in more detail here.

A table "translating" the notation used here to that used in the AMPL code is found on page 8.

The scenario model

The scenario tree used is illustrated in the figure below.
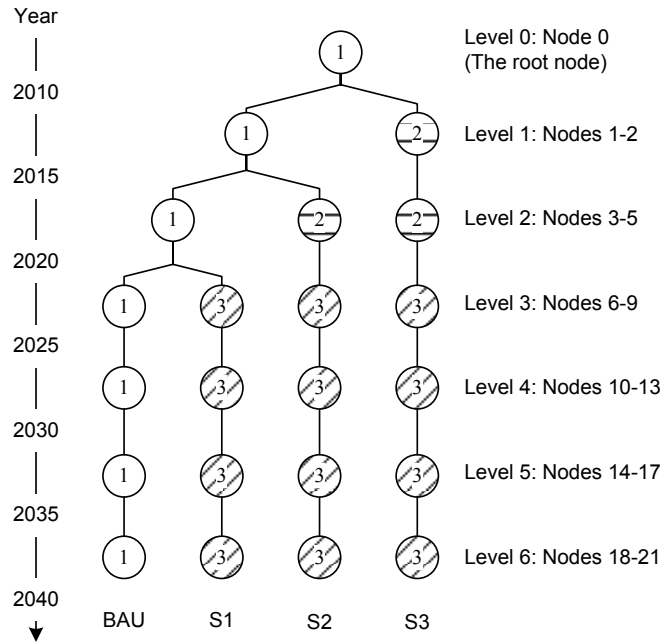
Every node n in the scenario tree corresponds to a specific realization of uncertain parameters at a specific stage. The scenario tree is here constructed from three building blocks, each representing one realization of the uncertain parameters. Node $n = 0$ is called the root, denoted by R, and corresponds to the known deterministic state at the beginning of the process. Each new level $\ell(n)$ of nodes n in the scenario tree represents a new stage in the decision process. A scenario is the realization of a root–to–leaf path in the tree. By construction, any node n except the root node has exactly one parent node $p(n)$ at the previous level. The set of all nodes n is denoted by N.

Data for the building blocks are found in the AMPL data file. In the AMPL file the scenarios are defined by the nodes of which they consist, and for each node, the corresponding building block is defined. The probability of each scenario is defined; this also determines the probability for each of the nodes. To limit the computation times for this project exercise, it is assumed that investments can only be made in the first three levels[1] of the scenario tree (i.e. for $\ell(n) \in \{1, 2, 3\}$).

The two objectives

The economic objective is to find the combination of investments resulting in the highest expected net present value (NPV). The objective is thus given by

$$\text{maximize} \quad f_{\text{NPV}} := \sum_{n \in N} \text{pr}^n \left[\phi(\ell(n)) f_R(\alpha^n, \xi^n) - \psi(\ell(n)) f_C(\hat{x}^n, \hat{y}^n, \delta^n) \right],$$

where pr^n is the probability of node n, ϕ and ψ are factors used to account for discount rates and time, f_R is the yearly revenue—a function of the energy exports α^n and the uncertain parameters ξ^n of node n—and f_C is the total capital expenditure—a function of the investment decisions, $\hat{x}^n$, $\hat{y}^n$, and δ^n, in node n.

The yearly revenue is calculated as the sum of energy exports multiplied by their net income, that is, the difference between the selling price and the operating cost. The total capital expenditure is given by the sum of the costs of the investments made in the node. A distinction is made in the model between fixed-cost measures $m \in M$, which (usually) result in a reduced use of steam at the mill, and investments $u \in U$ which make use of the acquired steam surplus to generate an energy export opportunity[2]. The cost of the first type of investment $m \in M$ is a constant. For the second type of investment $u \in U$ the cost is a function of the equipment size.

The CO_2 objective is to maximize the expected net CO_2 emissions reductions, and is given by

$$\text{maximize} \quad f_{\text{CO2}} := \sum_{n \in N} \text{pr}^n \phi(\ell(n)) f_{\text{Em}}(\alpha^n, \xi^n),$$

where f_{Em} is the yearly emissions reduction, which is a function of the energy exports α^n and the uncertain parameters ξ^n. According to the above formulation, the CO_2 emissions are discounted analogously as cash flows.

Here, the *ε-constraint method* will be used to solve the multiobjective optimization problem. In this method, only one of the objective functions is optimized for, while the other is reformulated as a constraint. Here, f_{NPV} is selected for optimization. The CO_2 objective is then reformulated as $f_{\text{CO2}} \geq \varepsilon$.

---

[1] In the original setting, described in the referenced article, investments could also be made at level 4.

[2] The set of "steam-using" options is given by $U = \{$BP, CT, LIG, DH100, DHLP, DH60$\}$ where BP=back-pressure turbine, CT=condensing turbine, LIG=lignin extraction, DH100=district heating using 100°C heat, DHLP=district heating using low-pressure steam, and DH60=district heating using 60°C heat.

The different preferences of the decision-makers are articulated by the different choices of the value of the limit ε.

An overview of variables and constraints

The binary variables $\hat{x}_m^n$ and $\hat{y}_u^n$ control the investment cost and take the value 1 in the node where investments are made. When an investment has been made, the investment becomes "active" in the following nodes. An active investment in steam savings means that the steam surplus is available for use, and an active investment in, e.g., a turbine, means that the turbine capacity is available. The binary variables x_m^n and y_u^n take the value 1 when the investment is "active". This relation between $\hat{x}_m^n$ and x_m^n is stated below. Corresponding constraints are valid for $\hat{y}_u^n$ and y_u^n.

$$x_m^0 = 0, \qquad\qquad\qquad m \in M, \qquad\qquad\qquad (1a)$$

$$x_m^n = x_m^{p(n)} + \hat{x}_m^{p(n)} - \check{x}_m^{p(n)}, \qquad\qquad m \in M,\ n \in N\backslash R. \qquad (1b)$$

(The variable $\check{x}_m^n$—and similarly $\check{y}_u^n$—is introduced for the possibility to deactivate investments.)

The steam savings from the measures $m \in M$ are given by the parameters s_{mp}. A steam balance on each steam pressure level $p \in P$ means that the steam flow ρ_{up}^n used for different options $u \in U$ must not exceed the acquired steam surplus plus the steam that is passed from higher pressure levels, minus the steam that is passed to lower pressure levels. As an example, the steam balance constraint for medium-pressure (MP) steam is given by

$$\sum_{u \in U\backslash(Q \cup L)} \rho_{u,\text{MP}}^n \leq \sum_{m \in M} x_m^n s_{m,\text{MP}} + \sum_{u \in Q} \left(\rho_{u,\text{HP}}^n - \rho_{u,\text{MP}}^n\right) h_{u,\text{MP}}, \quad n \in N. \quad (2)$$

Here, the set Q is a set of options, which use steam of one pressure and let it out at a lower pressure. The parameter h_{up} (in the inequality (2), $p = \text{MP}$) is a factor which is introduced due to the fact that when steam is passed from higher to lower pressures, it will be superheated, and water will be added to saturate the steam. L is a set of export options for which the above constraints are replaced by other similar ones due to yearly demand variations.

Another important constraint controls the amount of steam required to achieve a specific output α_u^n of power, heat, or lignin, and is given by

$$\alpha_u^n \leq \sum_{p \in P} q_{up}\rho_{up}^n + \gamma_u^n, \qquad\qquad u \in U\backslash L,\ n \in N. \qquad (3)$$

As described above, the variable ρ_{up}^n denotes the flow of steam with pressure p used for technology u in node n. The parameter q_{up} relates the steam of pressure p used for the power, lignin, or heat output for export option $u \in U$. (The variable γ_u^n is a possible additional output for option u in node n, which can be achieved without any steam input.)

The output α_u^n of power, lignin, and district heating for the different technology options is of course limited by the installed capacity β_u^n in node n for option u.

The relation between β_u^n and the variable δ_{ui}^n, which is used in the linearized cost function f_C, can be stated similarly to the constraints above, which relate x_m^n to $\hat{x}_m^n$. This is expressed as

$$\beta_u^0 = 0, \tag{4a}$$

$$\beta_u^n = \beta_u^{p(n)} + \sum_{i \in I_u} \delta_{ui}^{p(n)}, \qquad n \in N \backslash R, u \in U, \tag{4b}$$

where I_u is the set of linearization intervals used for export option $u \in U$. The capacity constraint is then given by

$$\alpha_u^n \leq \beta_u^n - g_u y_u^n, \qquad u \in U \backslash L, n \in N. \tag{5}$$

(The parameter g_u possesses a non-zero value only for the back-pressure turbine, for which it defines the current capacity of an existing turbine at the mill. By this formulation, the output α_u^n will represent a net increase of the electricity production in the turbine compared to today's level.)

Exercises

The optimization model is implemented in AMPL in the model file `co2.mod` with original data in `co2.dat`. The *basic case* assumed in these original model and data files is defined by

- a *single objective*: the maximization of f_{NPV}, and

- a *discrete uniform probability distribution* over the scenarios.

Use these original files and make necessary changes for solving the exercises.

- Choose objective function f_{NPV} or f_{CO2} (add/remove commenting).

- The constraint $f_{\text{CO2}} \geq \varepsilon$ is represented in AMPL by `CO2Decrease`. Relax this constraint by choosing lower values of ε (in AMPL: `co2_limit`), e.g. $\varepsilon = 0$.

Basics

1. Find the economically optimal investment plan, i.e. solve the model with no constraint on CO_2 emissions.

    (a) What is the expected net present value for this solution?

    (b) What is the resulting CO_2 reduction for this solution? (Hint: `display CO2Decrease.slack` yields the difference $f_{\text{CO2}} - \varepsilon$.)

    (c) What are the characteristics of the investment plan, i.e. what investments are made and when? The total investment plan must not be described in detail. (Hint: Display the AMPL variables `activate` to see which investments are made when and display `size` to see the invested capacity of turbines etc.)

2. Find the investment plan that maximizes the CO_2 emissions reductions, i.e. solve the model with the objective function f_{CO2} instead of f_{NPV}.

   (a) What is the CO_2 reduction for this solution?

   (b) What are the characteristics of the investment plan?

Constructing the Pareto graph

The solutions from 1) and 2) correspond to the extreme points of the Pareto curve (f_{NPV}^* and f_{CO2}^*). This means that the resulting CO_2 reductions for any Pareto-optimal solution will lie between the two values obtained for the CO_2 reductions.

3. Construct a graph (e.g. in Matlab or Excel) showing a number of (fairly well spread) points on the Pareto curve: Define a set of 5–15 values of ε between the minimum and maximum values established in tasks 1) and 2), and solve the model, maximizing f_{NPV} for each value in the set. Notice that, since the model is mixed-binary, there may be a positive slack in the CO_2 constraint also with increasing values of ε. The resulting CO_2 reduction should therefore be found considering this slack (i.e. f_{CO2} is in AMPL given by `CO2Decrease.slack + co2_limit`).

4. Discuss the appearance of the Pareto curve. Why is the decrease in f_{NPV} smaller when ε is increased for solutions close to the economic optimum?

Varying data – Sensitivity analysis

5. Investigate the sensitivity of the solution to variations in the assumed probability distribution:

   (a) Assume a probability of 1 for the BAU scenario, and probability 0 for the rest. Repeat tasks 1)–3) for this probability distribution. Discuss the differences and similarities between the different results.

   (b) Assume a probability of 1 for the S2 scenario, and probability 0 for the rest. Again repeat tasks 1)–3) for this new probability distribution. Discuss the results, and compare to those of 5(a).

6. Investigate the sensitivity of the solution to changes in various investment costs.

   (a) The cost for the recovery boiler upgrade (RBU) is difficult to determine, since the changes needed are different for each pulp mill. The value used in the model ($b_{RBU} = 29.8$ M€) was considered an upper estimate when the original study was carried out. The lower estimate for this cost was $b_{RBU} = 23.2$ M€. Change this value in the AMPL data file (the parameter name in AMPL is `measure_cost[RBU]`), and investigate how this affects the economic optimum for a discrete uniform probability distribution.

(b) The cost for district heating piping is assumed to be split between the pulp and paper company owning the pulp mill and the energy company owning the district heating network. Assume that the pulp and paper company must pay the entire cost, i.e. double this value (in AMPL: `measure_cost[Piping]`) and investigate how this affects the economic optimum for a discrete uniform probability distribution.

AMPL notation

Sets	
N	`NODES`
R	`ROOT`
P	`PRESSURE`
L	`VARIED`
M	`MEASURE`
U	`USE`
Q	`REDUC`
I_u	`1..ints[u]`
Variables	
α_u^n	`output[u,n]`
$\hat{x}_m^n,\ \hat{y}_u^n$	`activate[m,n], activate[u,n]`
$x_m^n,\ y_u^n$	`active[m,n], active[u,n]`
$\check{x}_m^n,\ \check{y}_u^n$	`deactivate[m,n], deactivate[u,n]`
δ_{ui}^n	`newsize[u,i,n]`
ρ_{up}^n	`flow[u,p,n]`
γ_u^n	`free[u,n]`
Parameters	
pr^n	`prob[n]`
ε	`co2_limit`
b_m	`measure_cost[m]`
q_{up}	`conv[u,p]`
s_{mp}	`steam[m,p]`
h_{up}	`quench[u,p]`
Other	
n	`n`
$\ell(n)$	`level[n]`
$p(n)$	`parent[n]`
$\phi(\ell(n))$	`pvf[level[n]-1]*pvf_sum`
$\psi(\ell(n))$	`pvf[level[n]]*(1-residual[level[n]])`
Constraints	
(1a-b)	`Activation`
(2)	`SteamBalance_MP`
(3)	`SteamToOutput`
(4a-b)	`InstalledSize`
(5)	`Capacity`