Römisch, W.; Schultz, R.

Decomposition of a Multi-Stage Stochastic Program for Power Dispatch

We develop a multi-stage stochastic program for the optimal dispatch of electric power under uncertain demand in a generation system comprising thermal and pumped storage hydro plants. Based on an abstract duality argument we propose an iterative decomposition scheme involving a non-smooth convex master problem and decoupled single-unit multi-stage stochastic programs.

1. Introduction

This paper deals with a multi-stage stochastic program for finding a cost-optimal dispatch of electric power in a power system comprising thermal power plants and pumped storage plants, which is typical for the eastern part of Germany. Stochasticity enters via the electric power demand, which is random and unveiled only in the course of the power generation process which covers a time horizon of up to 168 hours (one week). Hence, a scheme of alternate decision and observation underlies the dispatch of electric power: Fix the schedule (dispatch) for the first time interval, observe the demand for the second interval, fix the schedule for the second interval and so on. Under the assumption that the power demand of the first interval is known the schedule for that interval is deterministic and the remaining schedules are random. The multi-stage stochastic program elaborated below aims at finding an optimal schedule for the first interval given the operational constraints for the power system and a proper modelling of the stochastic power demand.

The present paper widens the scope adopted in a former article ([4]), where a two-stage stochastic planning model with simplified dynamics between decision and observation was studied. Further related work is contained in [16] where multi-stage stochastic programs for the unit commitment problem are analysed. The latter includes start-up and shut-down decisions of units into the optimization. Whereas our model allows duality statements, the (non-convex) unit commitment model leads to a duality gap, which, however, is getting smaller if the number of units in the model is increasing. Here, we focus on the simpler power dispatch model, elaborate duality and decomposition, but do keep in mind relations to the more general case.

2. Model

Let T denote the number of time intervals in the optimization horizon and $\{d_t : t = 1, \dots, T\}$ be the stochastic process (on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$) reflecting the randomness of power demand. We assume that the information on the power demand is complete for $t = 1$ and that it decreases with growing t. This is modelled by a nested sequence (filtration) of $\sigma-$fields

$$\mathcal{A}_1 = \{\emptyset, \Omega\} \subseteq \mathcal{A}_2 \subseteq \dots \subseteq \mathcal{A}_t \subseteq \dots \subseteq \mathcal{A}_T \subseteq \mathcal{A}$$

where d_t is $\mathcal{A}_t-$measurable $(t = 1, \dots, T\}$. (In particular, d_1 is then deterministic.)

Let $I_t \subseteq \{1, ..., I\}$ denote the index set of thermal units committed (i.e. on-line) at time t, with I denoting the number of available thermal units, and let J denote the number of pumped storage plants which are assumed to be on-line all the time, since both in the pumping and generation modes they can be driven upward from zero continuously. According to the stochasticity of power demand the scheduling decisions for the units are discrete-time stochastic processes as well:

$$\{p_t : t = 1, \dots, T\}, \quad \{(s_t, w_t) : t = 1, \dots, T\}.$$

Here, p_t^i $(i \in I_t,\ t = 1, \dots, T)$ denotes the output of the thermal unit i at time t and s_t^j, w_t^j $(j = 1, ..., J,\ t = 1, \dots, T)$ are the generation and pumping levels, respectively, of the pumped storage plant j in time step t. The following box constraints reflect output limitations of the units

$$p_{min}^i \le p_t^i \ \le \ p_{max}^i, \ \ 0 \le s_t^j \ \le \ s_{max}^j, \ \ 0 \le w_t^j \le w_{max}^j, \quad i \in I_t,\ j = 1, \dots, J,\ t = 1, \dots, T, \tag{1}$$

where $p_{min}^i, p_{max}^i, s_{max}^j, w_{max}^j$ are (non-stochastic) constants. Further operational constraints model availability restrictions and water balances in the pumped storage plants (for details see [4], [5])

$$\begin{array}{rcccl} S^j_{in} - S^j_{max} & \leq & \sum_{t=1}^{\tau}(s^j_t - \eta_j w^j_t) & \leq & S^j_{in}, \qquad j = 1,\ldots,J,\ \tau = 1,\ldots,T, \\ & & \sum_{t=1}^{T}(s^j_t - \eta_j w^j_t) & = & S^j_{Lev}, \qquad j = 1,\ldots,J. \end{array} \tag{2}$$

By η_j $(j = 1,\ldots,J)$ we denote the pumping efficiency which computes as the quotient of the energy gained when emptying the full upper dam and the energy needed when pumping upward the full content of the lower dam. We assume that there are no additional in- and outflows to the dams. The constants S^j_{in}, S^j_{max} denote the initial and maximal fill (in energy), respectively, of the upper dam, and S^j_{Lev} is a given fill of the upper dam at the end of the time horizon. The constraints (2) are crucial as they couple operation of units at different time steps. Thus, demand values at later times influence also actual decisions and the mentioned scheme of alternate decision and observation cannot be decoupled with respect to time. The following equations model the equilibrium between generation and (random) demand at all time steps

$$\sum_{i\in I_t} p^i_t + \sum_{j=1}^{J}(s^j_t - w^j_t) = d_t, \quad t = 1,\ldots,T. \tag{3}$$

These are the only constraints coupling operation of different units. A final constraint models the non-anticipativity of the stochastic process of decisions $\{(p_t, s_t, w_t) : t = 1,\ldots,T\}$. It says that, at time t, decisions (p_t, s_t, w_t) must not depend on future realizations of the process $\{d_t : t = 1,\ldots,T\}$. In other words, (p_t, s_t, w_t) has to be $\mathcal{A}_t-$measurable, which can be formalized by

$$(p_t, s_t, w_t) = \mathbf{E}((p_t, s_t, w_t)|\mathcal{A}_t), \quad t = 1,\ldots,T, \tag{4}$$

where $\mathbf{E}(\cdot|\mathcal{A}_t)$ denotes the conditional expectation with respect to $\mathcal{A}_t$. The objective function is given by the expected value of fuel costs of the thermal units

$$\mathbf{E}[\sum_{t=1}^{T}\sum_{i\in I_t} c^i(p^i_t)] = \mathbf{E}[\sum_{t=1}^{T}\sum_{i=1}^{I} a_{it}c^i(p^i_t)] \tag{5}$$

where $\mathbf{E}$ denotes expectation, $c^i(.)$ is a convex (linear, piecewise linear or quadratic) function and $a_{it} = 1$ for all $i \in I_t$ and $a_{it} = 0$ otherwise.

Altogether, (1) - (5) amounts to a multi-stage stochastic program which, via (3), is loosely coupled with respect to operation of different units. For larger power systems like the one considered here the number of stochastic variables in (1) - (5), which computes as $\sum_{t=1}^{T}$ card $(I_t) + 2JT$, is considerable. Numerical approaches are based on suitable discretizations of the demand distribution (scenario trees), which leads to large-scale optimization problems with usually millions of variables. In general, such problems are too large from the viewpoint of even the latest solution methods in multi-stage stochastic programming ([1], [2], [6], [13], [14]). Therefore, we present a decomposition scheme for (1) - (5) that employs solutions to smaller multi-stage stochastic programs for which existing solution methodology can be applied and adapted, respectively.

3. Duality and Decomposition

Let $x := \{x_t : t = 1,...,T\}$, with $x_t := (p_t,\, s_t,\, w_t)\,(t = 1,...,T)$, denote the decision process and $X := \times_{t=1}^{T} L_\infty(\Omega,\, \mathcal{A}_t, \mathcal{P}; \mathbb{R}^{n_t})$, with $n_t :=$ card $(I_t) + 2J(t = 1,...,T)$, the decision space equipped with the norm $\|x\| := \max_{t=1,...,T} \|x_t\|_\infty$. The fixed constraints (1), (2), (4) for the decision process are formalized by $C := \{x \in X : x(\omega) \in B, \mathcal{P} -$ a.s. $\}$ where B denotes the bounded polyhedron in $\mathbb{R}^m$ $(m := \sum_{t=1}^{T} n_t)$ given by the operational constraints (1), (2).

Further, let $Y := \times_{t=1}^{T} L_\infty(\Omega,\, \mathcal{A}_t, \mathcal{P};\, \mathbb{R})$ denote the data space, Y^* its dual and define the mapping $A : X \longrightarrow Y$ by $[Ax]_t := \sum_{i\in I_t} p^i_t + \sum_{j=1}^{J}(s^j_t - w^j_t)$ for all $x \in X$ and $t \in \{1,...,T\}$. Assuming that the stochastic demand d belongs to Y, (1)-(5) is equivalent to the abstract minimization problem

$$(P) \qquad \min\ \{f(x) : x \in C,\ Ax = d\},$$

where $f(x)$ denotes the objective function given by (5). Of course, $f : X \longrightarrow \mathbb{R}$ is convex and continuous. Together with (P) we consider the perturbed problem

$$(P_\xi) \qquad \min\ \{f(x) : x \in C,\ Ax = d + \xi\} \quad (\xi \in Y)$$

and denote its marginal value by $\varphi(\xi)$. It is well known that convex duality results hinge upon the behaviour of $\varphi(\cdot)$ at $\xi = 0$ ([11], [17]). We will make use of the following duality statement.

Proposition.
Let $d \in Y$ and assume the regularity condition: There exists $\epsilon > 0$ such that $\{v \in B : Av = d + y\} \neq \emptyset \quad \mathcal{P}-$ a.s. for all $y \in \mathbb{R}^T$ with $\|y\| \leq \epsilon$. Then we have

$$\varphi(0) = \sup_{\lambda \in Y^*} \inf_{x \in C} \{f(x) + \lambda(Ax - d)\}. \tag{6}$$

Proof. Let $\xi \in Y$ be such that $\|\xi\| := \text{ess sup}_{\omega \in \Omega} \|\xi(\omega)\| \leq \epsilon$. The regularity condition implies $\{v \in B : Av = d + \xi\} \neq \emptyset \quad \mathcal{P}-$ a.s. By utilizing a measurable selection argument [12] one shows analogously to the proof of Theorem 3.1 in [3] that there exists an element $x \in C$ such that $Ax = d + \xi$. Hence d belongs to the interior of the set $A(C) \subset Y$. In the terminology of [10] this means that the system $\{x \in C : Ax = d\}$ is regular. Theorem 1 of [10] then implies that the constraint set $M(\xi) := \{x \in C : Ax = d + \xi\}$ of (P_ξ) has the Hausdorff Lipschitz property $d_H(M(\xi), M(0)) \leq L_M \|\xi\|$ for all $\xi \in Y$ with $\|\xi\|$ sufficiently small (with some constant $L_M > 0$ and d_H denoting the Hausdorff distance). This property together with the Lipschitz continuity of f on bounded sets leads to a Lipschitz property of φ at 0. Appealing to the convex duality theorem in Section 2.2.3 of [17] completes the proof.$\square$

In terms of the power dispatch model the regularity condition says that, in each step $t \in \{1, ..., T\}$, the commitment schedule for the (on-line) thermal units has to fulfill a capacity (or reserve) constraint for $\mathcal{P}$ - almost all realizations of the random demand d_t .

According to the above duality statement, we consider the dual (concave) maximization problem

$$\max\{D(\lambda) : \lambda \in Y^*\}, \tag{7}$$

where $D(\lambda) := \inf\{f(x) + \lambda(Ax - d) : x \in C\}$ $(\lambda \in Y^*)$ and $Y^* := \times_{t=1}^T L_\infty^*(\Omega, \mathcal{A}_t, \mathcal{P}; \mathbb{R})$. For a general characterization of the duals to L_∞ the reader may consult e.g. Sect. 3 of [12]. Here we only use the observation that $Y^* := \times_{t=1}^T L_1(\Omega, \mathcal{A}_t, \mathcal{P}; \mathbb{R})$ holds as an isometry if $\mathcal{P}$ is a discrete probability measure with finite support. We confine ourselves to discrete $\mathcal{P}$ with finite support and focus on the decomposition structure of the dual function:

$$\begin{aligned} D(\lambda) &= \inf\{\mathbf{E}\left[\sum_{t=1}^T \sum_{i \in I_t} c^i(p_t^i)\right] + \sum_{t=1}^T \mathbf{E}[\lambda_t([Ax]_t - d_t)] : x \in C\} \\ &= \inf\{\sum_{i=1}^I \sum_{t=1}^T a_{it} \mathbf{E}\left[c^i(p_t^i) + \lambda_t p_t^i\right] + \sum_{j=1}^J \sum_{t=1}^T \mathbf{E}\left[\lambda_t(s_t^j - w_t^j)\right] - \sum_{t=1}^T \mathbf{E}[\lambda_t d_t] : (p,\, s,\, w) \in C\} \\ &= \sum_{i=1}^I \inf_{p^i} \left\{\sum_{t=1}^T a_{it} \mathbf{E}\left[c^i(p_t^i) + \lambda_t p_t^i\right]\right\} + \sum_{j=1}^J \inf_{(s^j, w^j)} \left\{\sum_{t=1}^T \mathbf{E}\left[\lambda_t(s_t^j - w_t^j)\right]\right\} - \sum_{t=1}^T \mathbf{E}[\lambda_t d_t] \end{aligned} \tag{8}$$

Here the infima are taken subject to the single-unit contraints (1), (2) for p^i and $(s^j,\, w^j)$.

Given $\lambda \in Y^*$, the evaluation of $D(\lambda)$ (and of a subgradient) requires the solution of $I + J$ multi-stage stochastic programs for all single (thermal, pumped storage hydro) units. Since the stochastic programs for the thermal units only contain box constraints (1), they can be solved explicitly for each $t \in \{1, ..., T\}$. The linear multi-stage model for each pumped storage plant contains only $2T$ stochastic variables and can be solved by existing solution methods (e.g. [1], [6], [14]).

The (iterative) decomposition approach for (1)-(5) now consists in solving (7) (with a discrete demand distribution) by convex nonsmooth minimization methods ([7], [8], [9], [15]) such as the bundle-trust method whose application is outlined next. Let $\bar{D}$ denote the convex function $-D$. An iteration step of the bundle method then looks as follows

$$\lambda_{k+1} = \lambda_k - \alpha_k \pi_k$$

where

$$\pi_k \in \text{argmin}\{\max_{i \in \Delta_k}\{g_i \pi - \bar{D}(\lambda_k) + \bar{D}(\mu_i) - g_i(\lambda_k - \mu_i)\} + \frac{1}{2\gamma_k}\|\pi\|^2 \ :\ \pi \in \mathbb{R}^N\}.$$

Here, λ_k denote the iteration points and μ_i are trial points that are accepted as iteration points if they fulfill certain descent conditions. Moreover, $g_i (i \in \Delta_k \subset \{1, \ldots, k\})$ belongs to the subdifferential $\partial \bar{D}(\mu_i)$ of $\bar{D}$ at μ_i, and $\alpha_k > 0, \gamma_k > 0$ are steplength and trust region parameters, respectively. For the model (7), the dimension N of the underlying Euclidean space computes as $N = \sum_{t=1}^{T}$ card supp d_t where card supp d_t denotes the cardinality of the support of the $t-$th component of the discrete random variable d.

Since $\bar{D}(\lambda) = \sup\{-f(x) + \lambda(d - Ax) \ : \ x \in C\}$, the subgradient formula for the maximum of convex functions in particular implies that g_i can be computed as $g_i = d - Ax(\mu_i)$ with $x(\mu_i) \in \operatorname{argmax}\{-f(x) + \mu_i(d - Ax) \ : \ x \in C\}$. Therefore, function values of D and subgradients of $-D$ that are needed for the solution of (7) can be computed by solving

$$\min\{f(x) + \lambda(Ax - d) \ : \ x \in C\}$$

for fixed values of λ. As mentioned above, the latter is accomplished by solving the single-unit multi-stage stochastic programs arising in (8).

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft.

4. References

1 Birge, J.R.; Donohue, C.J.; Holmes, D.E.; Svintsitski, O.G.: A parallel implementation of the nested decomposition algorithm for multistage stochastic linear programs, Technical Report 94-1, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor 1994.
2 Dupačová, J.: Multistage stochastic programs: The state-of-the-art and selected bibliography, Kybernetika 31(1995), 151-174.
3 Fiedler, O.; Römisch, W.: Stability in multistage stochastic programming, Annals of Operations Research 56(1995), 79-93.
4 Gröwe, N.; Römisch, W.; Schultz, R.: A simple recourse model for power dispatch under uncertain demand, Annals of Operations Research (to appear).
5 Guddat, J.; Römisch, W.; Schultz, R.: Some applications of mathematical programming techniques in optimal power dispatch, Computing 49(1992), 193–200.
6 Higle, J.L.; Sen, S.: A stochastic decomposition algorithm for multistage stochastic programs, personal communication, lecture presented at the 7th International Conference on Stochastic Programming, June 1995, Nahariya/Israel.
7 Kiwiel, K.C.: Proximity control in bundle methods for convex nondifferentiable minimization, Mathematical Programming 46(1990), 105-122.
8 Lemaréchal, C.: Lagrangian decomposition and nonsmooth optimization: Bundle algorithm, prox iteration, augmented Lagrangian, in: Nonsmooth Optimization Methods and Applications (F. Gianessi Ed.), Gordon and Breach, Amsterdam 1992, 201-216.
9 Mifflin, R.: A quasi-second-order proximal bundle algorithm, manuscript, Washington State University Pullman, 1993, revised 1995.
10 Robinson, S. M.: Stability theory for systems of inequalities. Part I: Linear Systems, SIAM Journal Numerical Analysis 12(1975), 754-769.
11 Rockafellar, R. T.: Duality and stability in extremum problems involving convex functions, Pacific Journal of Mathematics 21(1967), 167-187.
12 Rockafellar, R. T.: Integral functionals, normal integrands and measurable selections, in: Nonlinear Operators and the Calculus of Variations (G. P. Gossez et al. Eds.), Lecture Notes in Mathematics Vol. 543, Springer-Verlag, Berlin 1976, 157-207.
13 Rockafellar, R.T.; Wets, R. J-B: Scenarios and policy aggregation in optimization under uncertainty, Mathematics of Operations Research 16(1991), 119-147.
14 Rosa, C.; Ruszczyński, A.: On augmented Lagrangian decomposition methos for multistage stochastic programs, Working Paper WP-94-125, IIASA Laxenburg (Austria), 1994.
15 Schramm, H.; Zowe, J.: A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results, SIAM Journal on Optimization 2(1992),121-152.
16 Takriti, S.; Birge, J.R.; Long, E.: A stochastic model for the unit commitment problem, IEEE Transactions on Power Systems (to appear).
17 Tichomirov, V. M.: Grundprinzipien der Theorie der Extremalaufgaben. Teubner, Leipzig, 1982.

Addresses: Prof. Dr. Werner Römisch, Institut für Mathematik, Humboldt-Universität Berlin, Unter den Linden 6, D-10099 Berlin, Germany
Dr. Rüdiger Schultz, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Heilbronner Str. 10, D-10711 Berlin, Germany