Augmented Lagrangian Method for Multistage Stochastic Problems with Application to Energy Systems

Noel Smith
Advised by Darinka Dentcheva

Stevens Institute of Technology

May 8, 2019

Multistage Stochastic Problems

Formulation

$$\begin{aligned} \min_{x_1, \mathbf{x}_2, \dots, \mathbf{x}_T} \quad & \mathbb{E}\Big[f_1(x_1) + f_2(\mathbf{x}_2(\xi_{[2]}), \xi_2) + \cdots + f_T(\mathbf{x}_T(\xi_{[T]}), \xi_T) \Big] \\ \text{subject to} \quad & x_1 \in \mathcal{X}_1, \mathbf{x}_t(\xi_{[t]}) \in \mathcal{X}_t(\mathbf{x}_{t-1}(\xi_{[t-1]}), \xi_t), t = 2, \dots, T. \end{aligned} \tag{1}$$

where

- $\xi_t \in \mathbb{R}^d$ are uncertainty vectors revealed over T stages,
- $\xi_{[t]} := (\xi_1, \dots, \xi_t)$ is the history up to time t,
- $x_t = \mathbf{x}_t(\xi_{[t]}) \in \mathbb{R}^n$ are decision vectors considered functions of the uncertainty,
- $\mathcal{X}_t : \mathbb{R}^n \times \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ are measurable closed valued multifunctions.

Scenario Trees

Assume the data process ξ has finitely many realizations $\xi^k = (\xi_1^k, \ldots, \xi_T^k)$, each with probability $p_k > 0$, for $k = 1, \ldots, K$. Then the process can represented as a scenario tree with a unique root node representing the starting position. Denote:

- Ω_t the set of all nodes at stage $t = 1, \ldots, T$,
- for each $i \in \Omega_t$, the set of all child nodes $C_i \subset \Omega_{t+1}, t = 1, \ldots, T-1$
- $\Omega_{t+1} = \cup_{i \in \Omega_t} C_i$ and $C_i \cap C_j = \emptyset$ if $i \neq j$,
- $K = |\Omega_T|$.

Scenario Trees

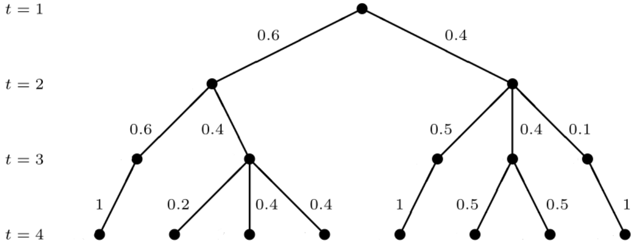


Figure: Example scenario tree with 4 stages and 8 scenarios. Nodes represent information states of the data process. Numbers along the arcs represent conditional probabilities of moving to the next node.

Polyhedral Multistage Problem

For each scenario ξ^k, there is a corresponding sequence of decisions $\mathbf{x}^k = (x_1^k, x_2^k, \dots, x_T^k)$. Consider polyhedral problems where the random objective functions $f_t(x_t, \xi_t)$ and constraint functions are polyhedral.

Polyhedral Formulation

$$
\begin{array}{lrcrcrcl}
\min & \sum_{k=1}^{K} p_k \sum_{t=1}^{T} f_t(x_t^k, \xi_t^k) & & & & & & \\
\text{s.t.} & A_{11}(\xi_1^k)x_1^k & & & & & = & b_1(\xi_1^k), \\
& A_{21}(\xi_2^k)x_1^k & + & A_{22}(\xi_2^k)x_2^k & & & = & b_2(\xi_2^k), \\
& & \ddots & & \ddots & & & \vdots \\
& & & A_{T,T-1}(\xi_T^k)x_{T-1}^k & + & A_{TT}(\xi_T^k)x_T^k & = & b_T(\xi_T^k), \\
& & & & & k = 1, \dots, K. & &
\end{array}
\tag{2}
$$

Nonanticipativity Constraints

At stage t, scenarios with the same history $\xi_{[t]}$ cannot be distinguished. This gives algebraic form for nonanticipativity constraints:

Nonanticipativity Constraints

$$x_t^k - x_t^l = 0, \text{ for all } k, l \text{ such that } \xi_{[t]}^k = \xi_{[t]}^l, \quad t = 1, \ldots, T. \tag{3}$$

A system of equations is formed by these constraints whose coefficients (I_n and $-I_n$) are defined by the block matrix $G = [G^1 \cdots G^K]$.
Then (3) can be written:

$$G^1 x^1 + \cdots + G^K x^K = 0.$$

Nonanticipativity Constraints

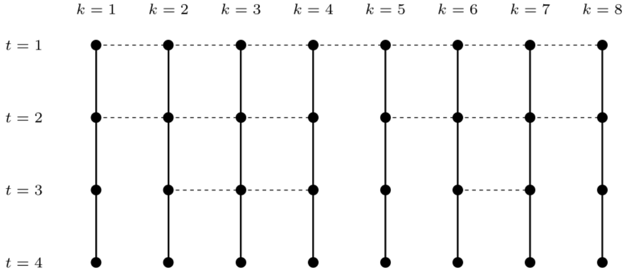


Figure: Sequences of decisions for scenarios from the example scenario tree. Horizontal lines represent equations of nonanticipativity

Formulation with Nonanticipativity Constraints

Objective function associated with scenario k

$$f^k(x^k) := \begin{cases} \sum_{t=1}^T f_t(x_t^k, \xi_t^k) & \text{if the constraints of (2) are satisfied for } k, \\ +\infty & \text{otherwise.} \end{cases}$$

Then formulation (2) can be written compactly and correctly:

$$\text{minimize} \left\{ f(x) := \sum_{k=1}^K p_k f^k(x^k) \right\} \quad \text{s.t.} \quad Gx = 0 \tag{4}$$

where $x = ((x^1)^\top, \ldots, (x^K)^\top) \in \mathbb{R}^{nKT}$.

Dual Decomposition

The Lagrangian associated with problem (4) is:

$$L(x, \lambda) := f(x) + \langle \lambda, Gx \rangle = \sum_{k=1}^{K} p_k f^k(x^k) + \sum_{k=1}^{K} \langle \lambda, G^k x^k \rangle.$$

The associated dual function has the form

$$D(\lambda) := \inf_x L(x, \lambda) = \sum_{k=1}^{K} \inf_{x^k \in \mathcal{X}^k} \{ p_k f^k(x^k) + \langle \lambda, G^k x^k \rangle \}$$

$$= \sum_{k=1}^{K} D^k(\lambda)$$

Dual Problem

$$\max_{\lambda \in \mathbb{R}} \sum_{k=1}^{K} D^k(\lambda)$$

Basic Augmented Lagrangian Method

The augmented Lagrangian associated with problem (4) has the form:

$$L_\rho(x, \lambda) := f(x) + \langle \lambda, Gx \rangle + \frac{\rho}{2}\|Gx\|^2, \tag{5}$$

where $\rho > 0$ is a penalty coefficient.

Augmented Lagrangian Method

Step 0 Set $r = 1$ and choose initial values for λ^1 and x^1.

Step 1 Calculate $x^r = \operatorname*{argmin}_x \ L_\rho(x, \lambda^r)$;

Step 2 If $Gx^r = 0$ then stop (optimal solution found). Otherwise, set $\lambda^{r+1} = \lambda^r + \rho G x^r$, increase r by one and return to Step 1.

The minimization step (1) cannot be easily decomposed into scenario subproblems due to the quadratic term $\|Gx\|^2$.

Distributed Augmented Lagrangian Method

An iterative nonlinear Jacobi method can be applied to overcome this non-separability.

Distributed Augmented Lagrangian for Multistage Problems

Step 0 Set $r = 1$ and choose initial values for λ^1 and x^1.

Step 1 For each scenario k solve the local simplified problem:

$$\hat{x}^{r,k} = \underset{x^k \in \mathcal{X}^k}{\operatorname{argmin}} \left\{ p_k f^k(x^k) + \langle \lambda^r, G^k x^k \rangle + \frac{\rho}{2} \left\| G^k x^k + \sum_{\sigma \neq k} G^\sigma x^{r,\sigma} \right\|^2 \right\}. \tag{6}$$

Step 2 With some stepsize $\tau \in (0, \frac{1}{2})$, update the primal variables

$$x^{r+1} = (1 - \tau) x^r + \tau \hat{x}^r.$$

Step 3 If $G x^{r+1} = 0$ and $G^k x^{r+1,k} = G^s \hat{x}^{r,k}$, the stop (the optimal solution is found).

Step 4 Set

$$\lambda^{r+1} = \lambda^r + \rho \tau G x^{r+1},$$

increase r by one and return to Step 1.

Convergence

Assumptions

(A1) The functions $f^k : \mathbb{R}^{nT} \to \mathbb{R},\ k = 1, \dots, K$, are convex and $\mathcal{X}^k \subseteq \mathbb{R}^{nT}$ are nonempty closed convex sets.

(A2) The Lagrange function L has a saddle point $(x^*, \lambda^*) \in \mathbb{R}^{nKT} \times \mathbb{R}$:

$$L(x^*, \lambda) \;\le\; L(x^*, \lambda^*) \;\le\; L(x, \lambda^*) \quad \forall\, x \in \mathcal{X},\ \lambda \in \mathbb{R}$$

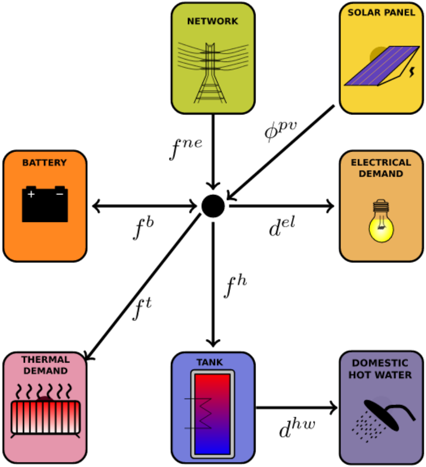
(A3) All scenario subproblems (6) are solvable at any iteration $r \in \mathbb{N}$

Energy Management System for Domestic Microgrid

Optimization problems in energy systems are prevaded with uncertainty and thus are apt for the application of stochastic programming models. To demonstrate the distributed augmented Lagrangian method, we introduce a model for the energy management system (EMS) of a domestic microgrid:

- Microgrids handle highly variable electric demands, heating demands, and renewable energy production
- Management forms a discrete time optimal control problem for time $t \in \{1, 2, \ldots, T = \frac{T_0}{\Delta}\}$
- A time horizon of $T_0 = 24$ hours with a time step of $\Delta = 15$ minutes is considered

Energy Management System for Domestic Microgrid



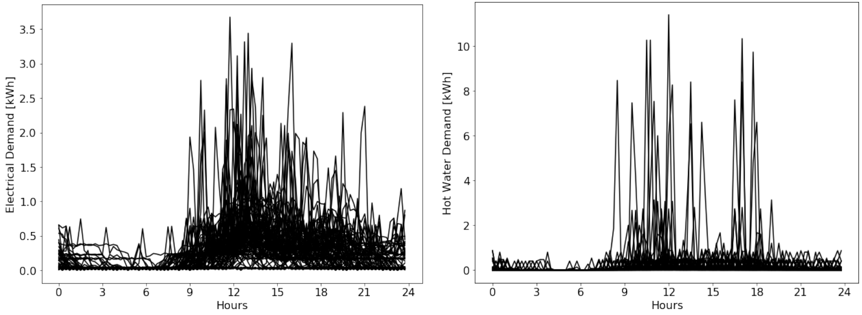
Load Balanace equation:

$$\phi^{pv}(t) + f^{ne}(t) = f^b(t) + f^t(t) + f^h(t) + d^{el}(t) \tag{7}$$

Uncertainties, Controls, and States

- **Uncertainties** $\mathbf{W}_t = (\mathbf{D}_t^{el}, \mathbf{D}_t^{th}, \mathbf{\Phi}_t^{pv})$
  - $\mathbf{D}_t^{el}$ electrical demand (kW)
  - $\mathbf{D}_t^{th}$ domestic hot water demand (kW)
  - $\mathbf{\Phi}_t^{pv}$ solar panel energy production (kW)
- **Control Variables** $\mathbf{U}_t = (\mathbf{F}_t^{b}, \mathbf{F}_t^{h}, \mathbf{F}_t^{t})$
  - $\mathbf{F}_t^{b}$ power to exchange with the battery (kW)
  - $\mathbf{F}_t^{t}$ power to heat hot water tank (kW)
  - $\mathbf{F}_t^{h}$ power used by electric heater (kW)
- **State Variables** $\mathbf{X}_t = (\mathbf{B}_t, \mathbf{H}_t, \theta_t^{w}, \theta_t^{i})$
  - $\mathbf{B}_t$ energy stored in the battery (kWh)
  - $\mathbf{H}_t$ energy stored in the hot water tank (kWh)
  - θ_t^{i} indoor temperature ($^\circ$C)
  - θ_t^{w} building wall temperature ($^\circ$C)

Some Scenarios of Uncertainties



(a) (b)

Figure: 100 scenarios for (a) electric and (b) hot water demands over the course of one day. The production of the solar panel $\mathbf{\Phi}^{pv}$ and electrical demands $\mathbf{D}^{el}$ are aggregated to only consider two uncertainites. Note the variability in these scenarios and their consistent shape, with peaks occuring at 12 pm and 8 pm and lulls at night. The spikes in hot water demands correspond to showers.

Discrete Time State Dynamics

Four linear state equations govern the evolution of the state variables over $t = 1, \ldots, T$:

$$
\begin{aligned}
\mathbf{B}_{t+1} &= \mathbf{B}_t + \Delta\big(\rho_c(\mathbf{F}^b_t)^+ - \frac{1}{\rho_d}(\mathbf{F}^b_t)^-\big)\\
\mathbf{H}_{t+1} &= \mathbf{H}_t + \Delta(\mathbf{F}^T_t - \mathbf{D}^{hw}_{t+1})\\
\boldsymbol{\theta}^w_{t+1} &= \boldsymbol{\theta}^w_t + \frac{\Delta}{c_m}\Big[\frac{\boldsymbol{\theta}^i_t - \boldsymbol{\theta}^w_t}{R_i + R_s} + \frac{\boldsymbol{\theta}^e_t - \boldsymbol{\theta}^w_t}{R_m + R_e} + \gamma\mathbf{F}^h_t + \frac{R_i}{R_i + R_s}\Phi^{int}_t + \frac{R_e}{R_e + R_m}\Phi^{ext}_t\Big]\\
\boldsymbol{\theta}^i_{t+1} &= \boldsymbol{\theta}^i_t + \frac{\Delta}{c_i}\Big[\frac{\boldsymbol{\theta}^w_t - \boldsymbol{\theta}^i_t}{R_i + R_s} + \frac{\boldsymbol{\theta}^e_t - \boldsymbol{\theta}^i_t}{R_v} + \frac{\boldsymbol{\theta}^e_t - \boldsymbol{\theta}^i_t}{R_f} + (1-\gamma)\mathbf{F}^h_t + \frac{R_i}{R_i + R_s}\Phi^{int}_t\Big]
\end{aligned}
$$

which is denoted

$$
\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) \quad \mathbb{P} - a.s. \tag{8}
$$

Optimization Criterion

- Cost to import electricity from the network

$$p_t^e \times \mathbf{F}_{t+1}^{ne}$$

where the recourse variable ensures the load balance equation is satisfied:

$$\mathbf{F}_{t+1}^{ne} = \mathbf{F}_t^b + \mathbf{F}_t^t + \mathbf{F}_t^h + \mathbf{D}_{t+1}^{el} - \mathbf{\Phi}_{t+1}^{pv} \quad \mathbb{P} - a.s.$$

- Virtual cost of thermal discomfort

$$p_t^d \times \max(0, \overline{\theta_t^i} - \boldsymbol{\theta}_t^i)$$

- The instantaneous piecewise linear costs are

$$L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) = p_t^e \times \mathbf{F}_{t+1}^{ne} + p_t^d \times \max(0, \overline{\theta_t^i} - \boldsymbol{\theta}_t^i)$$

- With a final linear cost

$$K(\mathbf{X}_T) = \kappa \times \max(0, \mathbf{X}_0 - \mathbf{X}_T)$$

Stochastic Optimization Problem Statement

EMS Problem

$$
\begin{aligned}
\min_{\mathbf{X},\mathbf{U}} \quad & \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T)\Big] \\
\text{s.t.} \quad & \mathbf{X}_0 = x_0, \\
& \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) \quad \mathbb{P}-a.s., \\
& \mathbf{U}_t \in \mathcal{U}_t(\mathbf{X}_t) \quad \mathbb{P}-a.s., \\
& \sigma(\mathbf{U}_t) \subset \mathcal{F}_t = \sigma(\mathbf{W}_1, \dots, \mathbf{W}_t).
\end{aligned}
\tag{9}
$$

where

- $\mathcal{U}_t$ is an admissible polyhedral set which set bounds on controls $\mathbf{U}_t$,
- $\sigma(\mathbf{U}_t) \subset \mathcal{F}_t$ is the nonanticipativity constraint.

Numerical Results for EMS Problem

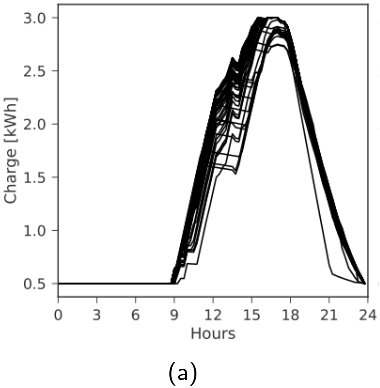


Figure: Energy stored in (a) battery and (b) hot water tank for a subset of 50 scenarios. In the numerical experiment, we considered a scenario tree with $K = 500$ scenarios and 6 time stages. Convergence (nonanticipativity constraint violation less than 10^{-3}) was found within 1800 iterations.