Nested Decomposition

Operations Research

Anthony Papavasiliou

Contents

Table of Contents

Scenario Trees

A **scenario tree** is a graphical representation of a Markov process $\{\xi_t\}_{t\in\mathbb{Z}}$, where

- nodes correspond to histories of realizations $\xi_{[t]} = (\xi_1, \ldots, \xi_t)$
- edges correspond to transitions from $\xi_{[t]}$ to $\xi_{[t+1]}$

Scenario Tree Terminology

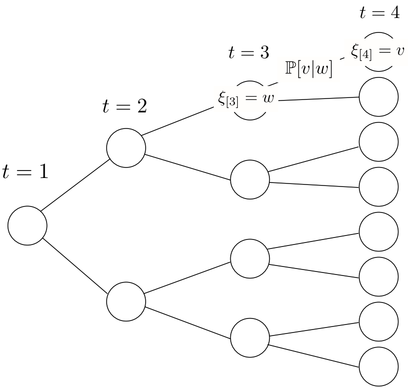
- Root corresponds to $t = 1$
- **Ancestor** of a node $\xi_{[t]}$, $A(\xi_{[t]})$: *unique* adjacent node which precedes ξ_t:

$$A(\xi_{[t]}) = \{\xi_{[t-1]} : (\xi_{[t-1]}, \xi_{[t]}) \in E\}$$

- **Children** or **descendants** of a node, $C(\xi_{[t]})$: set of nodes that are adjacent to $\xi_{[t]}$ and occur at stage $t+1$:

$$C(\xi_{[t]}) = \{\xi_{[t+1]} : (\xi_{[t]}, \xi_{[t+1]}) \in E\}$$

Scenario Tree Graphical Illustration



Specification of probability space requires:

- Assigning value $\xi_{[t]}$ for every node
- Assigning value $\mathbb{P}[\xi_{[t+1]}|\xi_{[t]}]$ for every edge

Recall that Ξ_t corresponds to the support of the random vector ξ_t, and $\Xi_{[t]} = \Xi_1 \times \ldots \times \Xi_t$ correponds to the support of the random process $\xi_{[t]} = (\xi_1, \ldots, \xi_t)$. In what follows, we will refer interchangeably to Ξ_t and S_t as the set of lattice nodes in stage t, and we will refer interchangeably to $S_1 \times \ldots \times S_t$ and $\Xi_{[t]}$ as the set of histories up to stage t. Furthermore, the notation $c_t(\omega_{[t]})$ and $c_{t,\omega_{[t]}}$ is used interchangeably for random variables, random vectors, and random matrices corresponding to node $\omega_{[t]}$ of a lattice.

Multi-Stage Stochastic Linear Programming on a Scenario Tree

$$(MSLP - ST):$$

$$\min c_1^T x_1 + \mathbb{E}[c_2(\omega_{[2]})^T x_2(\omega_{[2]}) + \cdots + c_H(\omega_H)^T x_H(\omega_{[H]})]$$

$$\text{s.t. } W_1 x_1 = h_1$$

$$T_1(\omega_{[2]})x_1 + W_2(\omega_{[2]})x_2(\omega_{[2]}) = h_2(\omega_{[2]}), \omega_{[2]} \in S_1 \times S_2$$

$$\vdots$$

$$T_{t-1}(\omega_{[t]})x_{t-1}(\omega_{[t-1]}) + W_t(\omega_{[t]})x_t(\omega_{[t]}) = h_t(\omega_{[t]}), \omega_{[t]} \in S_1 \times \ldots \times S_t$$

$$\vdots$$

$$T_{H-1}(\omega_{[H]})x_{H-1}(\omega_{[H-1]}) + W_H(\omega_{[H]})x_H(\omega_{[H]}) = h_H(\omega_{[H]}),$$

$$\omega_{[H]} \in S_1 \times \ldots \times S_H$$

$$x_1 \geq 0, x_t(\omega_{[t]}) \geq 0, t = 2, \ldots, H, \omega_{[t]} \in S_1 \times, \ldots, S_t$$

Notation

- Probability space $(\Omega, 2^{\Omega}, \mathbb{P})$
- We implicitly enforce **non-anticipativity** by requiring that x_t and ξ_t are functions of $\omega_{[t]}$
- $c_t(\omega) \in \mathbb{R}^{n_t}$: cost coefficients
- $h_t(\omega) \in \mathbb{R}^{m_t}$: right-hand side parameters
- $W_t(\omega) \in \mathbb{R}^{m_t \times n_t}$: coefficients of $x_t(\omega)$
- $T_{t-1}(\omega) \in \mathbb{R}^{m_t \times n_{t-1}}$: coefficients of $x_{t-1}(\omega)$
- $x_t(\omega)$: set of state *and* action variables in period t

Application of Dynamic Programming to Multi-Stage Stochastic Linear Programming

Step 1-a: Compute Q_H

$$
\begin{aligned}
Q_H(x_{H-1}, \xi_H) = \quad & \min_{x_H} c_H(\omega_{[H]})^T x_H \\
& \text{s.t. } T_{H-1}(\omega_{[H]}) x_{H-1} + W_H(\omega_{[H]}) x_H = h_H(\omega_{[H]}) \\
& x_H \geq 0
\end{aligned}
$$

Step 1-b: Compute V_H

$$V_H(x_{H-1}, \omega_{[H-1]}) = \mathbb{E}_{\xi_H}[Q_H(x_{H-1}, \xi_H) | \omega_{[H-1]}]$$

Recursive step a: Compute Q_t:

$$
\begin{aligned}
Q_t(x_{t-1}, \xi_t) = \quad & \min_{x_t} c_t(\omega_{[t]})^T x_t + V_{t+1}(x_t, \omega_{[t]}) \\
& \text{s.t. } T_{t-1}(\omega_{[t]}) x_{t-1} + W_t(\omega_{[t]}) x_t = h_t(\omega_{[t]}) \\
& x_t \geq 0
\end{aligned}
$$

Recursive step b: Compute V_t:

$$
V_t(x_{t-1}, \omega_{[t-1]}) = \mathbb{E}_{\xi_t}[Q_t(x_{t-1}, \xi_t) | \omega_{[t-1]}]. \tag{1}
$$

Final step: Solve for x_1:

$$
\begin{aligned}
& \min c_1^T x_1 + V_2(x_1) \\
& \text{s.t. } W_1 x_1 = h_1 \\
& x_1 \geq 0
\end{aligned}
$$

Structure of Value Function

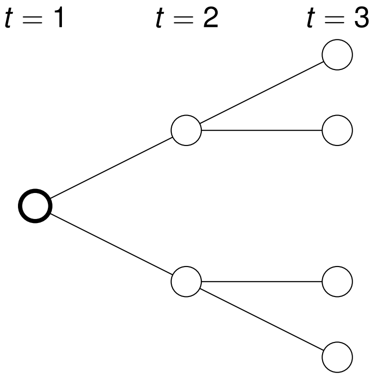
Consider a multi-stage stochastic linear program defined on a lattice, and denote $S_1 \times \ldots \times S_t$ as the set of nodes in stage t. If S_t is finite for all t then

- $V_{t+1,\omega_{[t]}}(x_t)$ and $Q_{t+1}(x_t, \xi_{t+1})$ are piecewise linear (pwl) convex
- dom $V_{t+1,\omega_{[t]}}$ and dom Q_{t+1} are polyhedral

Proof is by induction

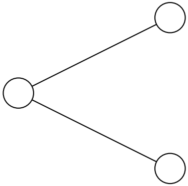
Table of Contents

Scenario Tree Model of Multi-Stage Stochastic Program



Goal: know what to do in the root node: $t = 1, k = 1$

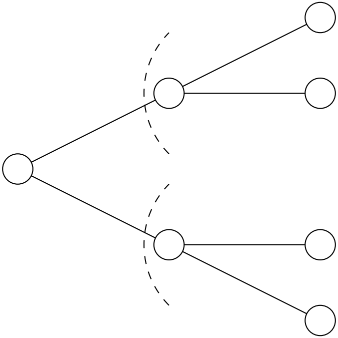
We know how to solve a 2-stage stochastic program



Algorithms

- L-shaped method
- Multi-cut L-shaped method

Breaking Down Multi-Stage to 2-Stage

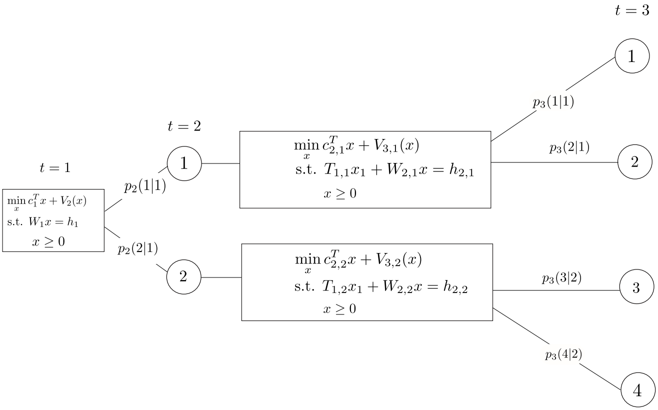


First index denotes time, second index denotes scenario

- Cost-to-go at $t = 2, k = 1$: piecewise linear function of $x_{2,1}$
- Cost-to-go at $t = 2, k = 2$: piecewise linear function of $x_{2,2}$
- Problem at $t = 1, k = 1$ has identical structure to 2-stage stochastic program

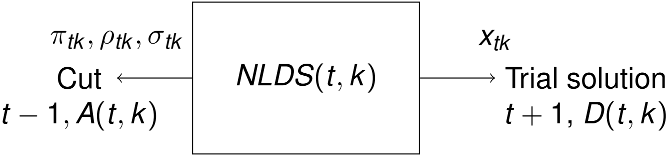
Idea of Nested Decomposition

- Each box corresponds to a linear program (why?)
- Nested decomposition: repeated application of the L-shaped method
- Variants depending on how we traverse the scenario tree



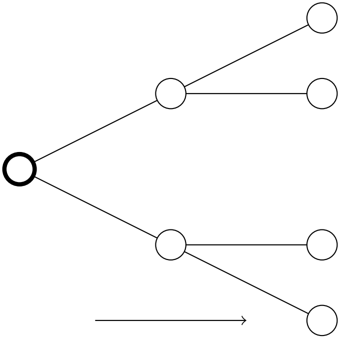
Nested L-Shaped Decomposition Subproblem (NLDS)

Building block: $NLDS(t,k)$: problem at stage t, scenario k



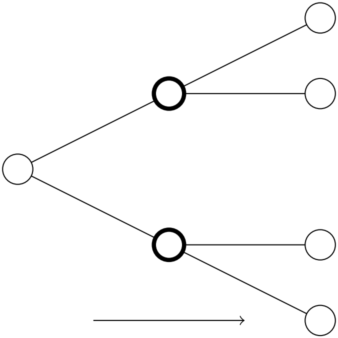
- $A(t,k)$: ancestor of outcome k in period t
- $D(t,k)$: descendants of outcome k in period t

Example



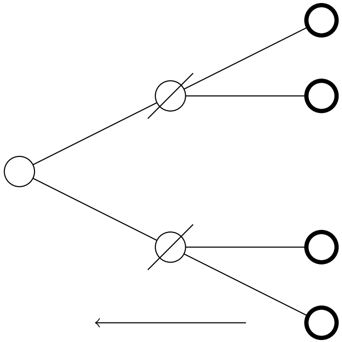
- Node: $(t = 1, k = 1)$
- Direction: forward
- Output: $x_{1,1}$

Example



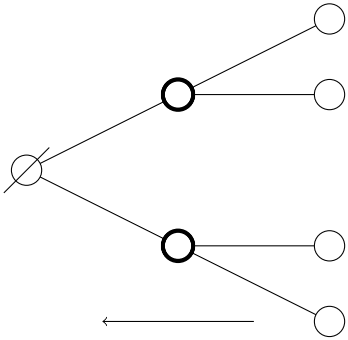
- Nodes: $(t = 2, k)$, $k \in \{1, 2\}$
- Direction: forward
- Output: $x_{2,k}$, $k \in \{1, 2\}$

Example



- Nodes: $(t = 3, k)$, $k \in \{1, 2, 3, 4\}$
- Direction: backward
- Output: $(\pi_{3,k}, \rho_{3,k}, \sigma_{3,k})$, $k \in \{1, 2, 3, 4\}$

Example



- Nodes: $(t = 2, k)$, $k \in \{1, 2\}$
- Direction: backward
- Output: $(\pi_{2,k}, \rho_{2,k}, \sigma_{2,k})$, $k \in \{1, 2\}$

Example: Newsboy Problem

Denote:

- C: unit cost of newspapers
- P: sales price of newspapers
- D_ω: random demand
- x: amount of newspapers procured (first stage)
- s: amount of papers sold (second stage)

Write out *NLDS* for stage 1 and 2

First stage:

$$
\begin{aligned}
& NLDS(1) : \min_x C \cdot x \\
& \text{s.t. } x \geq 0
\end{aligned}
$$

Second stage:

$$
\begin{aligned}
& NLDS(2, k) : \min_s -P \cdot s \\
& \text{s.t. } s \leq D_k \\
& s \leq x \\
& s \geq 0
\end{aligned}
$$

Example: Hydrothermal Scheduling

Denote

- C: marginal cost of thermal units
- E: reservoir capacity of hydroelectric dam
- $R_{t,k}$: rainfall (random)
- D_t: power demand
- x: hydro power stored in the dam
- q: hydro power production
- p: thermal production

Write out *NLDS* for stage t

NLDS for stage t and outcome k:

$$
\begin{aligned}
&NLDS(t,k) : \min_{x,q,p} C \cdot p \\
&\text{s.t. } x \leq E \\
&x \leq x_{t-1} + R_{t,k} - q \\
&p + q \geq D_t \\
&x, q, p \geq 0
\end{aligned}
$$

The Nested L-Shaped Decomposition Subproblem

For each stage $t = 1, \ldots, H-1$, scenario $k = 1, \ldots, |\Xi_{[t]}|$

$$
\begin{aligned}
NLDS(t,k): \quad & \min_{x,\theta} (c_{t,k})^T x + \theta \\
(\pi): \quad & W_{t,k} x = h_{t,k} - T_{t-1,k} x_{t-1,A(t,k)} \\
(\rho_j): \quad & E_{t,k,j} x + \theta \geq e_{t,k,j}, j = 1, \ldots, r_{t,k} \qquad (2) \\
(\sigma_j): \quad & D_{t,k,j} x \geq d_{t,k,j}, j = 1, \ldots, s_{t,k} \qquad (3) \\
& x \geq 0
\end{aligned}
$$

- $\Xi_{[t]}$: support of $\xi_{[t]}$
- $A(t,k)$: ancestor of realization k at stage t
- $x_{t-1,A(t,k)}$: current solution from $A(t,k)$
- Constraints (3): feasibility cuts
- Constraints (2): optimality cuts

Boundary Conditions

- For $t = 1$, $h_{t,k} - T_{t-1,k} x_{t-1,A(t,k)}$ is replaced by b
- For $t = H$, θ and constraints (2) and (3) are removed

Dual of $NLDS(t,k)$

$$\max_{\pi,\rho,\sigma} \pi^T(h_{t,k} - T_{t-1,k}x_{t-1,A(t,k)}) + \sum_{j=1}^{r_{t,k}} \rho_j^T e_{t,k} + \sum_{j=1}^{s_{t,k}} \sigma_j^T d_{t,k,j}$$

$$\text{s.t. } \pi^T W_{t,k} + \sum_{j=1}^{r_{t,k}} \rho_j^T E_{t,k,j} + \sum_{j=1}^{s_{t,k}} \sigma_j^T D_{t,k,j} \leq c_{t,k}^T$$

$$\sum_{j=1}^{r_{t,k}} 1^T \rho_j = 1$$

$$\rho_1, \ldots, \rho_{r_{t,k}} \geq 0$$

$$\sigma_1, \ldots, \sigma_{s_{t,k}} \geq 0$$

Feasibility Cuts

If $NLDS(t,k)$ is infeasible, solver returns $(\pi, \sigma_1, \ldots, \sigma_{s_{t,k}})$ with $\sigma_j \geq 0, j = 1, \ldots, s_{t,k}$, such that:

- $\pi^T(h_{t,k} - T_{t-1,k}x_{t-1,A(t,k)}) + \sum_{j=1}^{s_{t,k}} \sigma_j^T d_{t,k,j} > 0$
- $\pi^T W_{t,k} + \sum_{j=1}^{s_{t,k}} \sigma_j^T D_{t,k,j} \leq 0$

The following is a valid feasibility cut for $NLDS(t-1, a(k))$:

$$(FC) : D_{t-1,A(t,k)}x \leq d_{t-1,A(t,k)}$$

where

$$\begin{aligned} D_{t-1,A(t,k)} &= \pi^T T_{t-1,k} \\ d_{t-1,A(t,k)} &= \pi^T h_{tk} + \sum_{j=1}^{s_{t,k}} \sigma_j^T d_{t,k,j} \end{aligned}$$

Optimality Cuts

For all $k \in D_{t-1,j}$, solve $NLDS(t,k)$, then compute

$$E_{t-1,j} = \sum_{k \in D(t-1,j)} p_t(k|j) \cdot (\pi_{t,k})^T T_{t-1,k}$$

$$e_{t-1,j} = \sum_{k \in D(t-1,j)} p_t(k|j) \cdot ((\pi_{t,k})^T h_{t,k} +$$

$$\sum_{i=1}^{r_{t,k}} \rho_{t,k,i}^T e_{t,k,i} + \sum_{i=1}^{s_{t,k}} \sigma_{t,k,i}^T d_{t,k,i})$$

The following is an optimality cut for $NLDS(t-1,j)$:

$$E_{t-1,j} x + \theta \geq e_{t-1,j}$$

The Nested Decomposition Algorithm

Pass	t	k	Result	Action
F	1		Feasible	$t \leftarrow 2$, $k \leftarrow 1$, Store θ_1, x_1 Send x to $NLDS(2, j), j \in D(1)$
F	1		Infeasible	Infeasible, exit
F	$1 < t \leq H - 1$	$k < \lvert\Xi_{[t]}\rvert$	Feasible	$k \leftarrow k + 1$, Send x to $NLDS(t + 1, j), j \in D(t, k)$
F	$1 < t \leq H - 1$	$k < \lvert\Xi_{[t]}\rvert$	Infeasible	$k \leftarrow k + 1$ Add FC to $NLDS(t - 1, A(t, k))$
F	$1 < t \leq H - 1$	$\lvert\Xi_{[t]}\rvert$	Feasible	$t \leftarrow t + 1$, $k \leftarrow 1$ Send x to $NLDS(t + 1, j), j \in D(t, k)$ If $t = H - 1$ then Pass $\leftarrow$ B
F	$1 < t \leq H - 1$	$\lvert\Xi_{[t]}\rvert$	Infeasible	$t \leftarrow t + 1$, $k \leftarrow 1$ Add FC to $NLDS(t - 1, A(t, k))$ If $t = H - 1$ then Pass $\leftarrow$ B
B	$t \geq 2$	$k < \lvert\Xi_{[t]}\rvert$	Feasible	$k \leftarrow k + 1$, Store (π, ρ, σ)
B	$t \geq 2$	$k < \lvert\Xi_{[t]}\rvert$	Infeasible	$k \leftarrow k + 1$ Add FC to $NLDS(t - 1, A(t, k))$
B	2	$\lvert\Xi_{[t]}\rvert$	Feasible	Pass $\leftarrow$ F, $t \leftarrow 1$ Add OC to $NLDS(1)$ If $\theta_1 \geq e - Ex_1$: Optimal, exit
B	2	$\lvert\Xi_{[t]}\rvert$	Infeasible	Pass $\leftarrow$ F, $t \leftarrow 1$ Add FC to $NLDS(1)$
B	$t > 2$	$\lvert\Xi_{[t]}\rvert$	Feasible	$t \leftarrow t - 1$, $k \leftarrow 1$ Add OC to $NLDS(t - 1, A(t, k))$
B	$t > 2$	$\lvert\Xi_{[t]}\rvert$	Infeasible	$t \leftarrow t - 1$, $k \leftarrow 1$ Add FC to $NLDS(t - 1, A(t, k))$

Direction of Movement

Whenever $NLDS(t, k)$ is solved, the following data is generated

- If feasible:
  - Trial decision $x_{t,k}$ (can be sent forward)
  - Optimality cut (can be sent backwards)
- If infeasible: feasibility cut (can be sent backwards)

Alternative protocols

- **Fast-forward-fast-back**: move in current direction, as far as possible
- **Fast-forward**: move forward whenever possible
- **Fast-back**: move backwards whenever possible

If all $\Xi_{[t]}$ are finite sets and all x have finite upper bounds, then the nested L-shaped method converges finitely to an optimal solution

Proof: BL, page 268

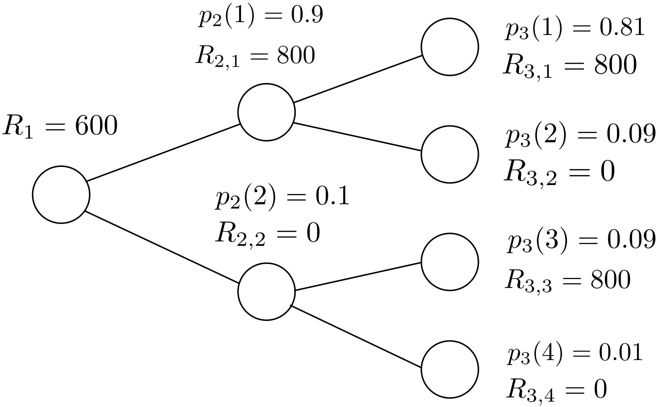
Table of Contents

Hydrothermal Scheduling over Three Periods

Consider the following hydrothermal problem:

- Demand: 1000 MW
- Energy capacity of dam: 750 MWh
- Marginal cost of thermal production: 25 $/MWh
- Capacity of thermal units: 500 MW
- Marginal cost of unserved demand: 1000 $/MWh

Scenario Tree



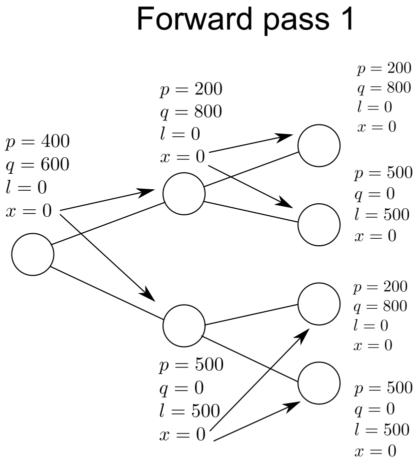
Is the tree serially independent?

NLDS

NLDS for first period:

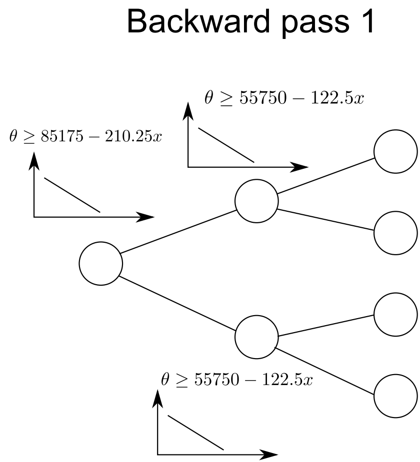
$$
\begin{aligned}
& NLDS(1) : \min 25 \cdot p + 1000 \cdot l \\
& \text{s.t. } x \leq 750 \\
& x \leq 600 - q \\
& p + q + l \geq 1000 \\
& p \leq 500 \\
& x, p, q, l \geq 0
\end{aligned}
$$

Algorithm Progress: Forward Pass 1



Greedy behavior $\Rightarrow$ load shedding in stage 2, node 2, and stage 3, nodes 2 and 4

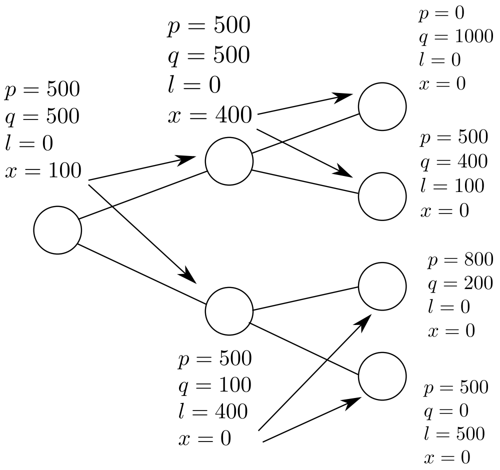
Algorithm Progress: Backward Pass 1



Cuts generated in stage 2 are identical (why?)

Algorithm Progress: Forward Pass 2

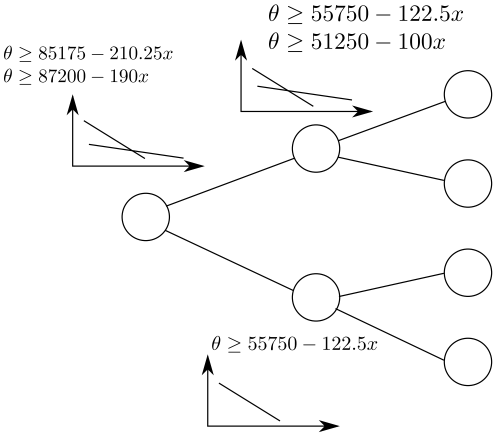
Forward pass 2



Note utilization of hydro in stage 1

Algorithm Progress: Backward Pass 2

Backward pass 2



New optimality cuts: node 1 of stage 2, stage 1

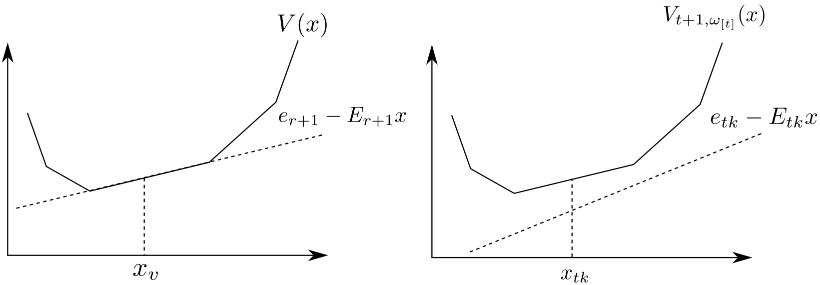
Convergence and Optimal Solution

Third forward pass $\rightarrow$ no new cut $\Rightarrow$ convergence

Load shedding in optimal policy: nodes 2 and 4 of stage 3

Optimal policy prevents spillage in scenarios of abundant water supply (node 1 of stage 3)

Optimality Cuts of L-Shaped Method and Nested Decomposition



- L-shaped method: optimality cuts support value function
- Nested decomposition: optimality cuts may be strictly below the value function