Stochastic Linear Programming

Operations Research

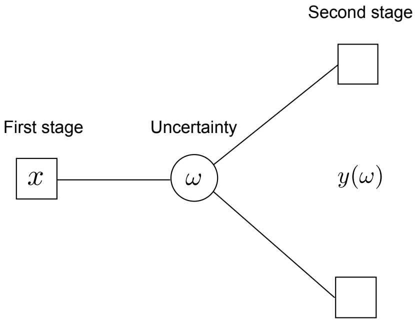
Anthony Papavasiliou

Stochastic Linear Programming

Outline

Sequence of Events

1. **First-stage decisions**: decisions taken before uncertainty is revealed
2. **Second-stage decisions**: decisions taken after uncertainty is revealed
3. Sequence of events: $x \to \omega \to y(\omega)$



Mathematical Formulation

$$\min c^T x + \mathbb{E}[\min q(\omega)^T y(\omega)]$$
$$Ax = b$$
$$T(\omega)x + W(\omega)y(\omega) = h(\omega)$$
$$x \geq 0, y(\omega) \geq 0$$

- First-stage decisions $x \in \mathbb{R}^{n_1}$, second stage decisions $y(\omega) \in \mathbb{R}^{n_2}$
- First-stage parameters: $c \in \mathbb{R}^{n_1}$, $b \in \mathbb{R}^{m_1}$, $A \in \mathbb{R}^{m_1 \times n_1}$
- Second-stage data: $q(\omega) \in \mathbb{R}^{n_2}$, $h(\omega) \in \mathbb{R}^{m_2}$, $T(\omega) \in \mathbb{R}^{m_2 \times n_1}$, $W(\omega) \in \mathbb{R}^{m_2, n_2}$
- **Fixed recourse** if W does not depend on ω

Example: Newsboy Problem

Denote

- x: amount of product produced in period 1
- y: amount of product sold in period 2
- C: unit cost of production
- P: sale price
- $D(\omega)$: random demand

Two-stage stochastic formulation of newsboy problem:

$$\begin{aligned} \min_{x, s(\omega) \geq 0} \; & C \cdot x - \mathbb{E}[P \cdot s(\omega)] \\ \text{s.t. } \; & s(\omega) \leq x \\ & s(\omega) \leq D(\omega) \end{aligned}$$

Extensions: salvage value, penalty for unserved demand

What is the trade-off of large/small value of x?

Example: Capacity Expansion Planning

$$\min_{x,y\geq 0} \sum_{i=1}^{n}(I_i \cdot x_i + \mathbb{E}[\sum_{j=1}^{m} C_i \cdot T_j \cdot y_{ij}(\omega)])$$

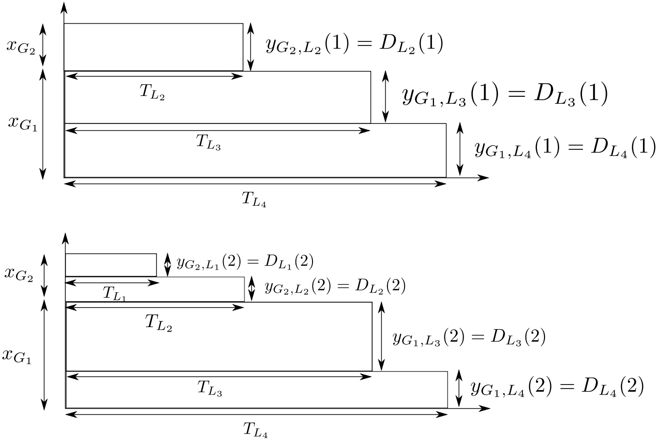
$$\text{s.t.} \sum_{i=1}^{n} y_{ij}(\omega) = D_j(\omega), j = 1, \ldots, m$$

$$\sum_{j=1}^{m} y_{ij}(\omega) \leq x_i, i = 1, \ldots n-1$$

- I_i, C_i: fixed/variable cost of technology i
- $D_j(\omega)$, T_j: height/width of load block j
- $y_{ij}(\omega)$: capacity of i allocated to j
- x_i: capacity of i

Note: T_j independent of ω

Example: Capacity Expansion Planning - Graphical Illustration



Note: T_j independent of ω

Example: Hydro-Thermal Scheduling

Denote:

- q_t: hydro power
- p_t: thermal power
- C: marginal cost of thermal power plant
- D_t: demand
- E: storage limit in the dam
- x_t: content of dam at the *end* of a stage
- r_t: amount of rain during stage t

Hydro-thermal scheduling problem:

$$\min C \cdot p_1 + \mathbb{E}[C \cdot p_2(\omega)]$$
$$p_1 + q_1 \geq D_1$$
$$x_1 \leq x_0 + r_1 - q_1$$
$$x_1 \leq E$$
$$p_2(\omega) + q_2(\omega) \geq D_2$$
$$q_2(\omega) \leq x_1 + r_2(\omega)$$
$$p_1, q_1, x_1, p_2(\omega), q_2(\omega) \geq 0$$

What is the trade-off?

Outline

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

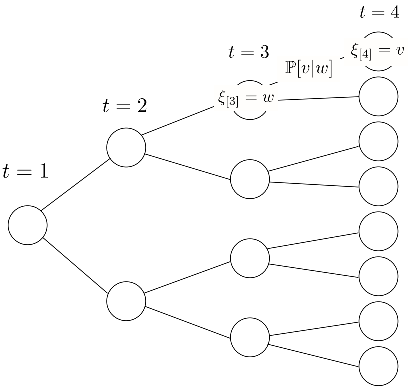
- Root corresponds to $t = 1$
- **Ancestor** of a node $\xi_{[t]}$, $A(\xi_{[t]})$: *unique* adjacent node which precedes ξ_t:

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

- **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:

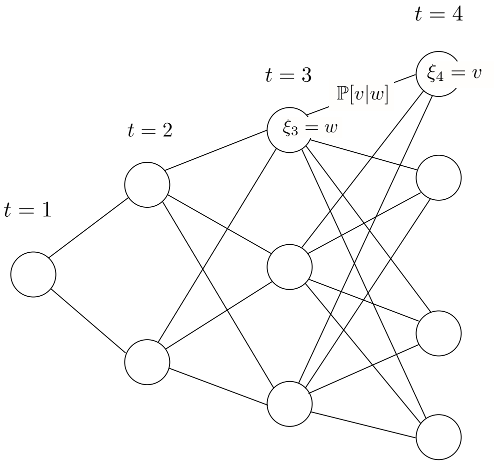
- Assigning value $\xi_{[t]}$ for every node
- Assigning value $\mathbb{P}[\xi_{[t+1]}|\xi_{[t]}]$ for every edge

Lattice

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

- nodes correspond to realizations ξ_t
- edges correspond to transitions from ξ_t to ξ_{t+1}

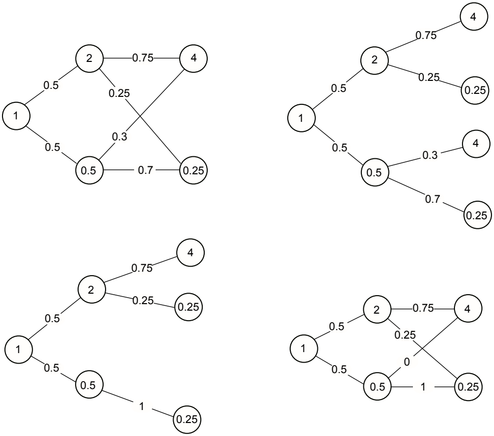
Lattice Graphical Illustration



Specification of probability space requires:

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

Equivalence of Scenario Trees and Lattices



We can

- unfold lattices into scenario trees (top)
- fold scenario trees into lattices (bottom)

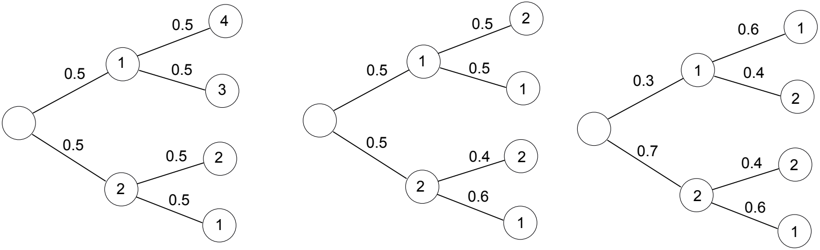
Serial Independence

A process satisfies **serial independence** if, for every stage t, ξ_t has a probability distribution that does not depend on the history of the process, i.e. one can define a probability measure $p_t(i)$ at each stage t, such that

$$\mathbb{P}[\xi_t(\omega) = i | \xi_{[t-1]}(\omega)] = p_t(i), \forall \xi_{[t-1]} \in \Xi_{[t-1]}, i \in \Xi_t$$

Checking for Serial Independence

Values on arcs indicate transition probabilities, values in nodes indicate realization of ξ_t



Which scenario tree(s) obey(s) serial independence

Populating Scenario Trees and Lattices with Data

- For scenario trees, one specifies:
  - The value of ξ_t at each node
  - The transition probability for every edge
- For lattices, one specifies:
  - The value of ξ_t at each node (a node generally does not correspond to a unique history $\xi_{[t]}$)
  - The transition probability for every edge
- For lattices with stage-wise independence, one specifies:
  - The value of ξ_t at each node
  - The probability of realization of each node of the lattice (well-defined)

Outline

General Formulation

Extended form of a multistage stochastic linear program:

$$
\begin{aligned}
&(MSLP): \\
&\min c_1^T x_1 + \mathbb{E}[c_2(\omega)^T x_2(\omega) + \cdots + c_H(\omega)^T x_H(\omega)] \\
&\text{s.t. } W_1 x_1 = h_1 \\
&T_1(\omega)x_1 + W_2(\omega)x_2(\omega) = h_2(\omega), \omega \in \Omega \\
&\vdots \\
&T_{t-1}(\omega)x_{t-1}(\omega) + W_t(\omega)x_t(\omega) = h_t(\omega), \omega \in \Omega \\
&\vdots \\
&T_{H-1}(\omega)x_{H-1}(\omega) + W_H(\omega)x_H(\omega) = h_H(\omega), \omega \in \Omega \\
&x_1 \geq 0, x_t(\omega) \geq 0, t = 2, \ldots, H
\end{aligned}
$$

Notation

- Probability space $(\Omega, 2^{\Omega}, \mathbb{P})$ with filtration $\{\mathcal{A}\}_{t \in \{1,\dots,H\}}$
- $c_t(\omega) \in \mathbb{R}^{n_t}$: 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
- We implicitly enforce **non-anticipativity** by requiring that x_t and ξ_t are *adapted* to filtration $\{\mathcal{A}\}_{t \in \{1,\dots,H\}}$

Stochastic Programs on Scenario Trees and Lattices

We now consider two specific instantiations of (MSLP):

- (MSLP-ST): stochastic programs on scenario trees
- (MSLP-L): stochastic programs on lattices

In these formulations, we will use the following notation:

- $\omega_t \in S_t$ (interpretation: index in the support Ξ_t of random input ξ_t)
- $\omega_{[t]} \in S_1 \times \ldots \times S_t$ (interpretation: index in $\Xi_{[t]} = \Xi_1 \times \ldots \times \Xi_t$, which is the history of realizations, up to period t)

Formulation on a Scenario Tree

$(MSLP - ST)$:

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

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

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

$$\vdots$$

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

$$\vdots$$

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

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

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

Formulation on Lattice

$$
\begin{aligned}
&(MSLP - L):\\
&\min c_1^T x_1 + \mathbb{E}[c_2(\omega_t)^T x_2(\omega_{[2]}) + \cdots + c_H(\omega_H)^T x_H(\omega_{[H]})]\\
&\text{s.t. } W_1 x_1 = h_1\\
&T_1(\omega_2)x_1 + W_2(\omega_2)x_2(\omega_{[2]}) = h_2(\omega_2), \omega_{[2]} \in S_1 \times S_2\\
&\vdots\\
&T_{H-1}(\omega_H)x_{H-1}(\omega_{[H]}) + W_H(\omega_H)x_H(\omega_{[H]}) = h_H(\omega_H),\\
&\omega_{[H]} \in S_1 \times \ldots \times S_H\\
&x_1 \geq 0, x_t(\omega_{[t]}) \geq 0, t = 2, \ldots, H
\end{aligned}
$$

Comments

- Compared to (MSLP-ST), ξ_t in (MSLP-L) is indexed over $\omega_t \in S_t$
- Problem size of (MSLP-L) doesn't really change compared to (MSLP-ST) (x_t is still indexed over $\omega_{[t]} \in S_1 \times \ldots \times S_t$)

Example: Capacity Expansion - Scenario Tree

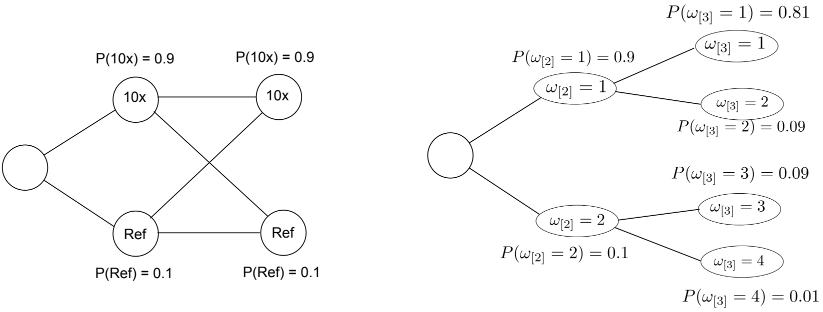


Table: Load duration curve for reference and 10x outcome

	Duration (hours)	Level (MW) Reference scenario	Level (MW) 10x wind scenario
Base load	8760	0-7086	0-3919
Medium load	7000	7086-9004	3919-7329
Peak load	1500	9004-11169	7329-10315

Example: Capacity Expansion - Technological Options

Technology	Fuel cost ($/MWh)	Inv cost ($/MWh)
Coal	25	16
Gas	80	5
Nuclear	6.5	32
Oil	160	2
DR	1000	0

Example: Capacity Expansion - Notation and Setup

Denote:

- $v_{it\omega_{[t]}}$: capacity of technology i constructed in period t
- $x_{it\omega_{[t]}}$: *total* amount of capacity of technology i available in period t
- $y_{ijt\omega_{[t]}}$: power allocation from technology i to load block j

Sequence of events:

1. Capacity $x_{i,t-1,\omega_{[t-1]}}$ available at the *end* of stage $t-1$ that can serve demand in t
2. Demand $D_{jt\omega_{[t]}}$ is observed
3. Construct new capacity $v_{it\omega_{[t]}}$

Example: Capacity Expansion - Model

Objective function:

$$\min_{x,v,y\geq 0} \sum_{i=1}^{n} I_i \cdot v_{i11}$$

$$+ \sum_{\omega_{[2]}=1}^{2} p_{\omega_{[2]}} (\sum_{i=1}^{n} I_i \cdot v_{i2\omega_{[2]}} + \sum_{i=1}^{n} \sum_{j=1}^{m} C_i \cdot T_j \cdot y_{ij2\omega_{[2]}})$$

$$+ \sum_{\omega_{[3]}=1}^{4} p_{\omega_{[3]}} (\sum_{i=1}^{n} I_i \cdot v_{i3\omega_{[3]}} + \sum_{i=1}^{n} \sum_{j=1}^{m} C_i \cdot T_j \cdot y_{ij3\omega_{[3]}})$$

Note: first stage involves *only* investment decision

Supply equals demand (enforced only for $t > 1$):

$$\sum_{i=1}^{n} y_{ijt\omega_{[t]}} = D_{jt\omega_{[t]}}, j \in \{1, \ldots, m\}, t \in \{2, \ldots, 3\}$$

$$\omega_{[2]} \in \{1, 2\}, \omega_{[3]} \in \{1, \ldots, 4\}$$

Example: Capacity Expansion - Model

Investment dynamics:

$$x_{i2\omega_{[2]}} = x_{i11} + v_{i2\omega_{[2]}}, i \in \{1, \ldots, n-1\}, \omega_{[2]} \in \{1, 2\}$$
$$x_{i3\omega_{[3]}} = x_{i21} + v_{i3\omega_{[3]}}, i \in \{1, \ldots, n-1\}, \omega_{[3]} \in \{1, 2\}$$
$$x_{i3\omega_{[3]}} = x_{i22} + v_{i3\omega_{[3]}}, i \in \{1, \ldots, n-1\}, \omega_{[3]} \in \{3, 4\}$$

Technology capacity constraints:

$$\sum_{j=1}^{m} y_{ij2\omega_{[2]}} \leq x_{i11}, i \in \{1, \ldots n-1\}, \omega_{[2]} \in \{1, 2\}$$
$$\sum_{j=1}^{m} y_{ij3\omega_{[3]}} \leq x_{i21}, i \in \{1, \ldots n-1\}, \omega_{[3]} \in \{1, 2\}$$
$$\sum_{j=1}^{m} y_{ij3\omega_{[3]}} \leq x_{i22}, i \in \{1, \ldots n-1\}, \omega_{[3]} \in \{3, 4\}$$

Does this model obey block separability?

Example: Capacity Expansion - Optimal Solution

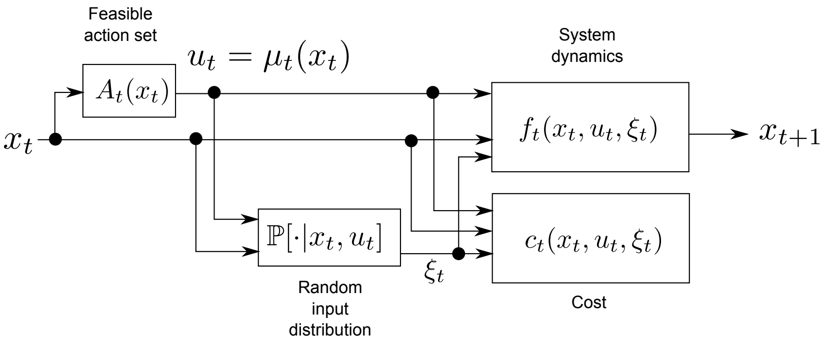
Optimal expansion plan:

- Coal, period 1: 2986 MW
- Nuclear, period 1: 7329 MW
- Oil, period 1: 854 MW
- Period 2: nothing (!)

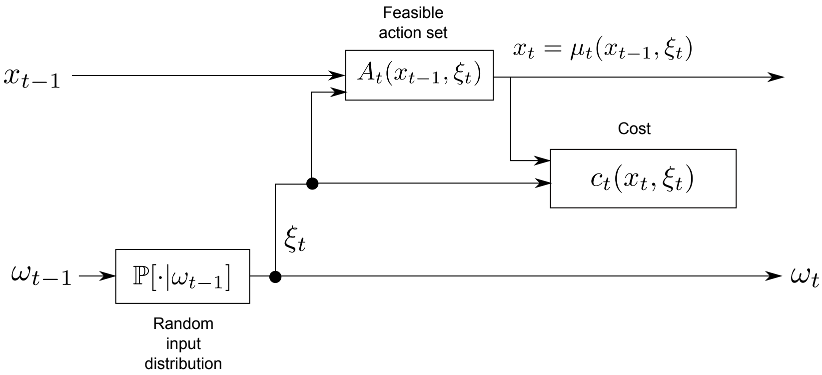
Why is it optimal to invest *only* in period 1?

Outline

Stochastic Control Block Diagram



Stochastic Programming Block Diagram



Comparison

- Timing of action
  - Stochastic control: *first* decide u_t, *then* observe realization of uncertainty ξ_t
  - Stochastic programming: *first* observe the realization of uncertainty, ξ_t, *then* decide x_t
- System state
  - Stochastic control: x_t encodes *all* information about system state
  - Stochastic programming: vector x_t *and* node of the lattice ω_t encode all information about system state
- Feasible action set A_t
  - Stochastic control: A_t depends only on x_t
  - Stochastic programming: A_t depends on x_{t-1} *and* ξ_t

Block Separability

Feasible action set in stage t:

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

Block separability occurs when these constraints can be written in the following form:

$$T^{xx}_{t-1}(\omega_t)x_{t-1}(\omega_{[t-1]}) + W^{xx}_t(\omega_t)x_t(\omega_{[t]}) = h^{xx}_t(\omega_t), \omega_{[t]} \in S_1 \times \ldots \times S_t$$
$$T^{xu}_{t-1}(\omega_t)x_{t-1}(\omega_{[t-1]}) + W^{xu}_t(\omega_t)u_t(\omega_t) = h^{xu}_t(\omega_t), \omega_{[t]} \in S_1 \times \ldots \times S_t$$

Benefit: decision variables u_t do not need to be propagated forward

Q-**function** in final period:

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

Value function in final period:

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

Application of Dynamic Programming in MSLP

Proceeding recursively, Q-function in stage t:

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

Value function in stage t:

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

Application of Dynamic Programming to MSLP

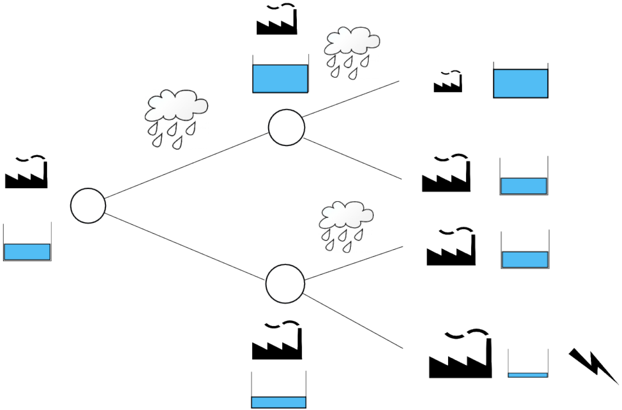
Proceed backwards until:

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

Notational Convention

- Note the different notation for node of the lattice (ω_t) and realization of uncertainty (ξ_t)
- The notation $V_t(x_{t-1}, \omega_{t-1})$ emphasizes how value functions are stored by SDDP
- The notation $Q_t(x_{t-1}, \xi_t)$ is the conventional notation used in stochastic programming, but Q-functions are not explicitly stored in SDDP
- Note the difference in the definition of the Q function
  - Stochastic control: function of state x, action u
  - Stochastic programming: function of state x, random input ξ_t

Example: Hydrothermal Scheduling



Example: Hydrothermal Scheduling

Consider the following hydro-thermal system:

- 3 periods
- Demand in each period: 1000 MW
- Marginal cost of thermal generators: 25 $/MWh
- Max production of thermal generators: 500 MW
- Marginal cost of lost load: 1000 $/MWh
- Rainfall: *independent* identically distributed, uniformly on $[0, 1000]$ MW, denote density function as $f : \mathbb{R} \to \mathbb{R}$

Example: Hydrothermal Scheduling

Denote

- p: thermal production
- q: hydro production
- l: unserved demand
- x_2: stored hydro energy at *beginning* of period 2

$$
\begin{aligned}
Q_3(x_2, R_3) = \quad & \min 1000 \cdot l + 25 \cdot p \\
& \text{s.t. } l + p + q \geq 1000 \\
& p \leq 500 \\
& q \leq x_2 + R_3 \\
& l, p, q \geq 0
\end{aligned}
$$

Q function of period 3:

$$Q_3(x_2, R_3) = \begin{cases} 0, x_2 + R_3(\omega) \geq 1000 \\ 25 \cdot (1000 - (x_2 + R_3(\omega))), 500 \leq x_2 + R_3(\omega) < 1000 \\ 500 \cdot 25 + 1000 \cdot (500 - (x_2 + R_3(\omega))), 0 \leq x_2 + R_3(\omega) < 500 \end{cases}$$

Value function of period 3:

$$\begin{aligned} & V_3(x_2) = \mathbb{E}_{R_3}[Q_3(x_2, R_3)] \\ & = \mathbb{P}[R_3(\omega) \geq 1000 - x_2] \cdot 0 \\ & + \int_{r=500-x_2}^{1000-x_2} (25 \cdot (1000 - r - x_2)) f(r) dr \\ & + \int_{r=0}^{500-x_2} (500 \cdot 25 + 1000 \cdot (500 - r - x_2)) f(r) dr \\ & = \begin{cases} 0, & x_2 \geq 1000 \\ 12500 - 25 \cdot x_2 + 0.0125 \cdot x_2^2, & 500 \leq x_2 < 1000 \\ 134375 - 512.5 \cdot x_2 + 0.5 \cdot x_2^2, & 0 \leq x_2 < 500 \end{cases} \end{aligned}$$

Note:

- V_3 is convex
- V_3 is *not* a piecewise linear function of x_2

Q_2 can be computed as:

$$
\begin{aligned}
&Q_2(x_1, R_2) = \min 1000 \cdot l + p + V_3(x_2) \\
&\text{s.t. } l + p + q \geq 1000, p \leq 500 \\
&x_2 = x_1 - q + R_2(\omega) \\
&l, p, q, x_2 \geq 0
\end{aligned}
$$

Q_2 yields V_2, V_2 yields Q_1, ...