IS for Jackson Networks with a Tree Topology

Devin Sezer

Middle East Technical University, Ankara Turkey

RESIM 2008, INRIA, Rennes, France

September 25, 2008

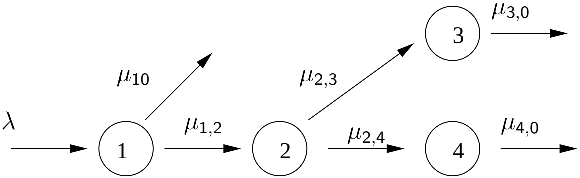


Figure: A tree network

Dynamics

X a random walk on $\mathbb{Z}_+^d$ with iid increments Y.

$$v_{i,j}(i) = -1, \quad v_{i,j}(j) = 1.$$

$$P(Y_k = v_{i,j}) = \mu_{i,j}, \quad P(Y_k = v_{0,1}) = \lambda.$$

$$X_{k+1} = X_k + \pi(X_k, Y_k).$$

Boundaries of the State Space

Refer to these boundaries by bitmaps $b \in \{0,1\}^d$. For example: $(0,0,\ldots,0)$ is the origin, i.e., all queues empty.

Buffer Structure

Two structures of interest:

$$\mathcal{S}_1 = \left\{ \sum_{i=1}^{d} x(i) = 1 \right\}$$

$$\mathcal{S}_2 = \{ x(i) = \beta(i) \text{ for some } i \text{ and } x(j) \leq \beta(j) \text{ for all } j \}, \quad \beta \in \mathbb{R}^d_+.$$

Stability

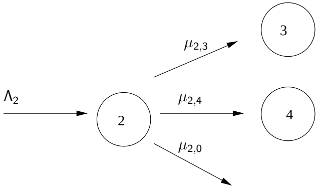
The utility of node *i*:

$$\rho_i \doteq \Lambda_i/\mu_i.$$

Stability

The utility of node i:

$$\rho_i \doteq \Lambda_i / \mu_i .$$



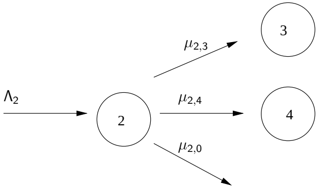
$$\mu_2 = \mu_{2,3} + \mu_{2,4} + \mu_{2,0}$$

$$\Lambda_3 = \Lambda_2 \frac{\mu_{2,3}}{\mu_2}$$

Stability

The utility of node i:

$$\rho_i \doteq \Lambda_i/\mu_i.$$



$$\mu_2 = \mu_{2,3} + \mu_{2,4} + \mu_{2,0}$$

$$\Lambda_3 = \Lambda_2 \frac{\mu_{2,3}}{\mu_2}$$

Assume

$$\bigvee_{i=1}^{d} \rho_i < 1.$$

Goal

Probability of Interest:

$$P_s(X_n \text{ hits } S_n \text{ before } 0).$$

$s = (1, 0, 0, .., 0)$ or any other point close to the origin.

To estimate this using efficient Importance Sampling one has to solve:

$$V_n(s) \doteq \inf_{\bar{P}} \mathbb{E}_s \left[1_{A_n}(X) \frac{dP}{d\bar{P}} \right].$$

Goal cont.

$$\lim -\frac{1}{n} \log V_n(s) = 2\gamma$$

holds.

Goal cont.

$$\lim -\frac{1}{n} \log V_n(s) = 2\gamma$$

holds.

Instead of exact solution of the optimization problem , find $\bar{P}_n$ such that:

$$\underline{\lim} -\frac{1}{n} \log \mathbb{E}_s \left[1_{A_n}(X) \frac{dP}{d\bar{P}_n} \right] \geq 2\gamma.$$

Such an IS change of measure is said to be **asymptotically optimal.**

The goal is to build such changes of measure.

References

Basic references on the subsolution approach to IS are [1, 2, 3, 4].

Dynamic Programming For V_n

$$V_n(x) = \inf_{\bar{\mu}_{i,j}} \sum V_n(x + \pi(x, v_{i,j}))\frac{\bar{\mu}_{i,j}}{\mu_{i,j}}.$$

Dynamic Programming For V_n

$$V_n(x) = \inf_{\bar{\mu}_{i,j}} \sum V_n(x + \pi(x, v_{i,j})) \frac{\bar{\mu}_{i,j}}{\mu_{i,j}}.$$

Boundary condition:

$$V_n(x) = 1, x \in S_n.$$

Dynamic Programming For V_n

$$V_n(x) = \inf_{\bar{\mu}_{i,j}} \sum V_n(x + \pi(x, v_{i,j})) \frac{\bar{\mu}_{i,j}}{\mu_{i,j}}.$$

Boundary condition:

$$V_n(x) = 1, x \in S_n.$$

V_n decays exponentially. For a meaningful asymptotic analysis define

$$W_n = -\frac{1}{n} \log V_n.$$

Dynamic Programming for W_n

Relative entropy representation:

$$-\log \int e^{-f} d\alpha = \inf_{\beta} \{ \int f d\beta + R(\beta | \alpha) \}.$$

Dynamic Programming for W_n

Relative entropy representation:

$$-\log \int e^{-f} d\alpha = \inf_{\beta} \{ \int f d\beta + R(\beta|\alpha) \}.$$

Apply this to the DPE in the previous slide to get:

$$W_n(x) = \sup_{\bar{\mu}_{i,j}} \inf_{\hat{\mu}_{i,j}} \sum \left(W_n(x + \pi(x, v_{i,j})) + \log \frac{\bar{\mu}_{i,j}}{\mu_{i,j}} + \log \frac{\tilde{\mu}_{i,j}}{\mu_{i,j}} \right) \hat{\mu}_{i,j}.$$

Dynamic Programming for W_n

Relative entropy representation:

$$-\log \int e^{-f} d\alpha = \inf_{\beta} \{ \int f d\beta + R(\beta|\alpha) \}.$$

Apply this to the DPE in the previous slide to get:

$$W_n(x) = \sup_{\bar{\mu}_{i,j}} \inf_{\hat{\mu}_{i,j}} \sum \left(W_n(x + \pi(x, v_{i,j})) + \log \frac{\bar{\mu}_{i,j}}{\mu_{i,j}} + \log \frac{\tilde{\mu}_{i,j}}{\mu_{i,j}} \right) \hat{\mu}_{i,j}.$$

Possible to switch the inf and the sup:

$$W_n(x) = \inf_{\bar{\mu}_{i,j}} \sum \left(W_n(x + \pi(x, v_{i,j})) + 2 \log \frac{\bar{\mu}_{i,j}}{\mu_{i,j}} \right) \bar{\mu}_{i,j}.$$

The limit PDE and boundary condition:

$$H_{b_x}(DW(x)) = 0, \qquad W(x) = 0 \text{ for } x \in \mathcal{S},$$

where

$$N_b(q) \doteq \lambda e^{-q(1)/2} + \sum_{i:b(i)=1} \mu_{i,j} e^{\frac{q(i)-q(j)}{2}} + \sum_{i:b(i)=0} \mu_i,$$

$$H_b(q) = -2 \log N_b(q),$$

$b_x \in \{0,1\}^d$ is the boundary on which x lies:

$$b_x(i) \doteq \begin{cases} 0, & \text{if } x(i) = 0, \\ 1, & \text{otherwise.} \end{cases}$$

Subsolutions to the limit PDE

$\bar{V}$ is an ϵ-subsolution to the limit equation if it is $C^1(\mathbb{R}^d, \mathbb{R})$ and

(a) $H_{b_x}(D\bar{V}(x)) \geq -\epsilon$ for all $x \in \mathbb{R}^d_+$

(b) $\bar{V}(0) - \bar{V}(x) \geq 2\gamma - \epsilon, x \in \mathcal{S}$.

Subsolutions to the limit PDE

$\bar{V}$ is an ϵ-subsolution to the limit equation if it is $C^1(\mathbb{R}^d, \mathbb{R})$ and

(a) $H_{b_x}(D\bar{V}(x)) \geq -\epsilon$ for all $x \in \mathbb{R}^d_+$

(b) $\bar{V}(0) - \bar{V}(x) \geq 2\gamma - \epsilon, x \in \mathcal{S}$.

How to generate IS change of measure from a subsolution W:

$$\bar{\mu}_{i,j} = \mu_{i,j} \exp(0.5(DW(x/n)_i - DW(x/n)_j) - H(DW(x/n))$$

if the random walk X is at the point $x \in \mathbb{Z}^d_+$.

IS Change of Measure given by the subsolution

$$\bar{\mu}_{i,j} = \mu_{i,j} \, \frac{\rho_i}{\rho_j},$$

renormalize if necessary.

IS Change of Measure given by the subsolution

$$\bar{\mu}_{i,j} = \mu_{i,j} \, \frac{\rho_i}{\rho_j},$$

renormalize if necessary.
ρ_i, ρ_j are $\rho_i(b_x)$ and $\rho_j(b_x)$, i.e., they depend on which boundary X is.

IS Change of Measure given by the subsolution

$$\bar{\mu}_{i,j} = \mu_{i,j} \, \frac{\rho_i}{\rho_j},$$

renormalize if necessary.
ρ_i, ρ_j are $\rho_i(b_x)$ and $\rho_j(b_x)$, i.e., they depend on which boundary X is.

$$\rho_i(b) = \frac{\Lambda_i}{M_i(b)}.$$

where

$$M_i(b) \doteq \begin{cases} \mu_i, & \text{if } b_i = 1 \\ \min\left(\mu_i, \sum_{k:i\to k} M_k(b) + \Lambda_i \frac{\mu_{i,0}}{\mu_i}\right), & \text{if } b_i = 0. \end{cases}$$

Corresponding subsolution

The effective gradient $q = (q_1, q_2, \ldots, q_d)$ **associated with bitmap** b**:**

$$q_i \doteq 2 \log \frac{\Lambda_i}{M_i(b)}.$$

Corresponding subsolution

The effective gradient $q = (q_1, q_2, \ldots, q_d)$ **associated with bitmap** b**:**

$$q_i \doteq 2 \log \frac{\Lambda_i}{M_i(b)}.$$

$$W_l^{c,\epsilon}(x) = -\alpha_l \epsilon + \langle q_l, x \rangle,$$

$$W^{c,\epsilon}(x) = \bigwedge_{l=1}^{L} W_l^{c,\epsilon}(x)$$

Corresponding subsolution

The effective gradient $q = (q_1, q_2, \ldots, q_d)$ **associated with bitmap** b**:**

$$q_i \doteq 2 \log \frac{\Lambda_i}{M_i(b)}.$$

$$W_l^{c,\epsilon}(x) = -\alpha_l \epsilon + \langle q_l, x \rangle,$$

$$W^{c,\epsilon}(x) = \bigwedge_{l=1}^{L} W_l^{c,\epsilon}(x)$$

One smoothing method that is simple and easy to implement on a computer is the following [3]. Define

$$W^{\epsilon,\delta}(x) \doteq -\delta \log \sum_{l=1}^{L} \exp \left\{ -\frac{1}{\delta} W_k^{c,\epsilon}(x) \right\}.$$

Simulation Results

In all IS Simulations 10000 samples.

Simulation Results

In all IS Simulations 10000 samples.

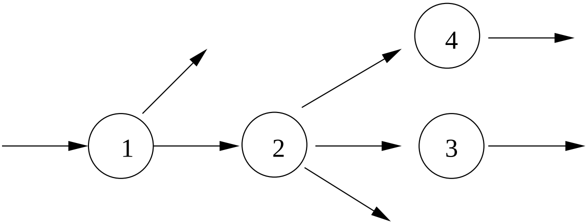


Figure: Example 1

$$\lambda = 0.01, \mu_{1,2} = \mu_{1,0} = 0.27,$$
$$\mu_{2,0} = \mu_{2,3} = \mu_{2,4} = 0.09,$$
$$\mu_{3,0} = \mu_{3,1} = \mu_{4,0} = \mu_{4,1} = 0.09.$$

Simulation Results Cont.

Exact probability $p_{30} = 2.749 \times 10^{-47}$.

	Estimate $\hat{p}_n$	Standard Error	95 % CI
Est. 1	2.96×10^{-47}	0.13×10^{-47}	$[2.69, 3.23] \times 10^{-47}$
Est. 2	3.00×10^{-47}	0.15×10^{-47}	$[2.69, 3.31] \times 10^{-47}$
Est. 3	2.74×10^{-47}	0.12×10^{-47}	$[2.50, 2.97] \times 10^{-47}$
Est. 4	2.75×10^{-47}	0.12×10^{-47}	$[2.52, 2.99] \times 10^{-47}$
Est. 5	2.72×10^{-47}	0.12×10^{-47}	$[2.47, 2.96] \times 10^{-47}$
Est. 6	2.83×10^{-47}	0.14×10^{-47}	$[2.55, 3.11] \times 10^{-47}$
Est. 7	2.95×10^{-47}	0.14×10^{-47}	$[2.67, 3.22] \times 10^{-47}$
Est. 8	2.57×10^{-47}	0.11×10^{-47}	$[2.35, 2.79] \times 10^{-47}$
Est. 9	2.72×10^{-47}	0.12×10^{-47}	$[2.48, 2.96] \times 10^{-47}$
Est. 10	2.75×10^{-47}	0.12×10^{-47}	$[2.51, 2.98] \times 10^{-47}$

Table: Simulation Result for Example 1

Simulation Results Cont.

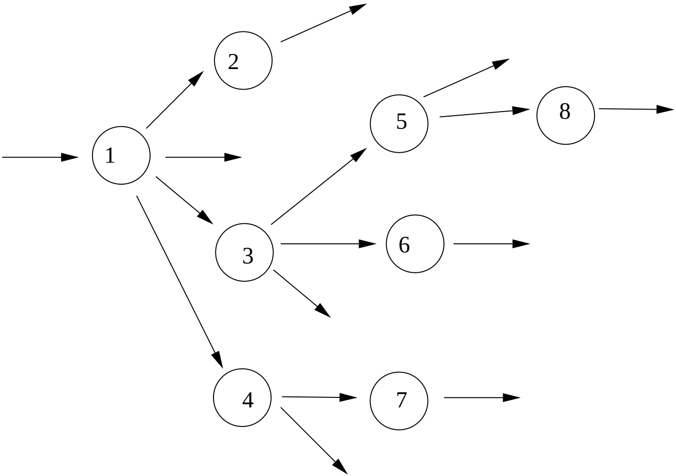


Figure: Example 2

Simulation Results Cont.

We take the arrival rate $\lambda = 0.1248$, The service rates at each node are

$$\mu_1 = 0.3748, \quad \mu_2 = 0.06, \quad \mu_3 = 0.18, \quad \mu_4 = 0.06$$
$$\mu_5 = 0.073, \quad \mu_6 = 0.073, \quad \mu_7 = 0.025, \quad \mu_8 = 0.028.$$

The routing matrix is:

$$R = \begin{pmatrix} 0 & 0.1666 & 0.5 & 0.1666 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.4 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Using this matrix one computes $\mu_{i,j}$ as

$$\mu_{i,j} = \begin{cases} \mu_i R_{i,j}, & j > 0 \\ \mu_i - \sum_{k=1}^{8} \mu_{i,k}, & j = 0. \end{cases}$$

Simulation Results Cont.

A straightforward simulation with 10^8 samples estimate p_{30} to be 1.2×10^{-6} with a standard error of 1.1×10^{-6}.

Simulation Results Cont.

	Estimate $\hat{p}_n$	Standard Error	95 % CI
Est. 1	1.11×10^{-6}	0.17×10^{-6}	$[0.78, 1.44] \times 10^{-6}$
Est. 2	1.69×10^{-6}	0.32×10^{-6}	$[1.04, 2.34] \times 10^{-6}$
Est. 3	1.25×10^{-6}	0.18×10^{-6}	$[0.89, 1.61] \times 10^{-6}$
Est. 4	1.94×10^{-6}	0.51×10^{-6}	$[0.92, 2.97] \times 10^{-6}$
Est. 5	1.23×10^{-6}	0.17×10^{-6}	$[0.89, 1.56] \times 10^{-6}$
Est. 6	1.67×10^{-6}	0.27×10^{-6}	$[1.13, 2.21] \times 10^{-6}$
Est. 7	1.77×10^{-6}	0.45×10^{-6}	$[0.87, 2.68] \times 10^{-6}$
Est. 8	1.38×10^{-6}	0.38×10^{-6}	$[0.61, 2.15] \times 10^{-6}$
Est. 9	1.30×10^{-6}	0.18×10^{-6}	$[0.94, 1.65] \times 10^{-6}$
Est. 10	1.25×10^{-6}	0.15×10^{-6}	$[0.94, 1.55] \times 10^{-6}$

Table: Simulation results for the network with eight nodes

Seperate buffer for each queue

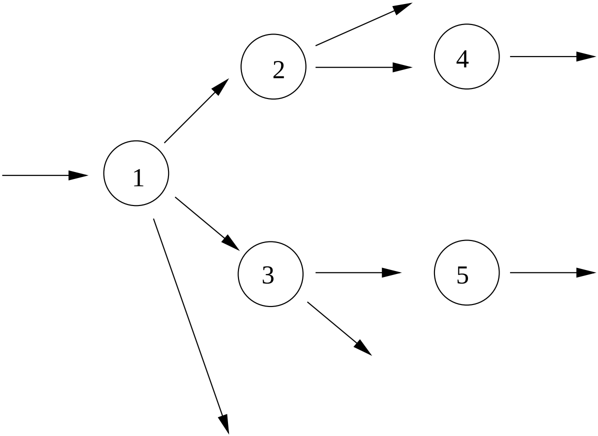


Figure: Example 3

Seperate buffer for each queue cont.

Routing matrix:

$$R = \begin{pmatrix} 0 & 0.2 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and the service and arrival rates are:

$$\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = 0.19, \quad \lambda = 0.05.$$

We will suppose that the buffer sizes for the nodes are respectively:

$$15, 15, 17, 18, 19.$$

Seperate buffer for each queue cont.

Exact probability $p_{19} = 6.8601 \times 10^{-9}$

	Estimate $\hat{p}_n$	Standard Error	95 % CI
Est. 1	7.33×10^{-9}	0.42×10^{-9}	$[6.50, 8.17] \times 10^{-9}$
Est. 2	6.81×10^{-9}	0.34×10^{-9}	$[6.12, 7.50] \times 10^{-9}$
Est. 3	7.30×10^{-9}	0.38×10^{-9}	$[6.53, 8.06] \times 10^{-9}$
Est. 4	7.05×10^{-9}	0.39×10^{-9}	$[6.28, 7.83] \times 10^{-9}$
Est. 5	7.01×10^{-9}	0.37×10^{-9}	$[6.26, 7.76] \times 10^{-9}$
Est. 6	6.31×10^{-9}	0.32×10^{-9}	$[5.67, 6.96] \times 10^{-9}$
Est. 7	6.78×10^{-9}	0.36×10^{-9}	$[6.06, 7.49] \times 10^{-9}$
Est. 8	7.03×10^{-9}	0.38×10^{-9}	$[6.27, 7.80] \times 10^{-9}$
Est. 9	6.92×10^{-9}	0.39×10^{-9}	$[6.14, 7.70] \times 10^{-9}$
Est. 10	7.24×10^{-9}	0.37×10^{-9}	$[6.51, 7.98] \times 10^{-9}$

Table: Simulation results for the case when each node has a separate buffer

The paper

The article on this work is at:

http://arxiv.org/abs/0708.3260

References

Paul Dupuis and Hui Wang.
Importance sampling, large deviations and differential games.
Stochastics and Stochastic Reports, 76(6):481–508, 2004.

Paul Dupuis and Hui Wang.
Adaptive importance sampling for uniformly recurrent markov chains.
Annals of Applied Probability, 15(1):1–38, 2005.

Paul Dupuis and Hui Wang.
Subsolutions of an isaacs equation and efficient schemes for importance sampling: Examples and numerics.
2005.
preprint available at `http://www.dam.brown.edu/lcds/publications`.

Paul Dupuis Ali Devin Sezer and Hui Wang.
Dynamic importance sampling for queueing networks.
Annals of Applied Probability, 17(4):1306–1346, 2007.

Thank you!