Efficient stochastic sensitivity analysis of discrete event systems

Paper by Sergey Plyasunov and Adam P. Arkin
Talk by Chris Baumann

June 7, 2013
why?
Sensitivities

Examples for outputs

- number of molecules of a certain type at a specific time
- when was the first time there were 1000 molecules of a certain type

In general:

Process at times $t_1 < t_2 < \ldots < t_M = T$

- output: $F(X(t_1), \ldots, X(t_M))$
- averaged output:
  
  $$o(k_1, k_2, \ldots) = \mathbb{E}[F(X(t_1), \ldots, X(t_M)) | \text{Parameters } k_1, k_2, \ldots]$$
Sensitivities

Examples for outputs

- number of molecules of a certain type at time $T$
- when was the first time there were 1000 molecules of a certain type

For simplicity:
look only at process at time $T$

- output: $f(T) := F(X(T))$
- averaged output:
  $$o(k_1, k_2, \ldots) = \mathbb{E}[f(T)|\text{Parameters } k_1, k_2, \ldots]$$
Sensitivities

Examples for outputs

- number of molecules of a certain type at time $T$
- when was the first time there were 1000 molecules of a certain type

For simplicity:
look only at process at time $T$

- output: $f(T) := F(X(T))$
- averaged output:
  \[ o(k) = \mathbb{E}[f(T)|\text{Parameter } k] \]
Derivatives

\[ o(k) \]

\[ k \]
Derivatives

\[
\frac{do}{dk} = \lim_{\epsilon \to 0} \frac{o(k(1 + \epsilon)) - o(k)}{\epsilon}
\]

simple, easily comparable
Approximation of derivatives

\[
\frac{d o}{d k} \approx \frac{o(k(1 + \epsilon)) - o(k)}{\epsilon}
\]

- how many simulation runs?
- which \( \epsilon \)?
Chemical reactions

\[ \nu_1^+ X_1 + \nu_2^+ X_2 + \ldots + \nu_S^+ X_S \xrightarrow{k} \nu_1^- X_1 + \nu_2^- X_2 + \ldots + \nu_S^- X_S \]

- reaction rate: \( a(X) = k \cdot h(X) \)
  - example: \( A + 2B \xrightarrow{k} C \) with \( h(A, B, C) = A \cdot B^2 \)

in general

- \( R \) reactions with rates \( a_r(X) = k_r \cdot h_r(X) \) for \( r = 1, \ldots, R \)
\[
\frac{do}{dk} = \lim_{\epsilon \to 0} \frac{o(k(1 + \epsilon)) - o(k)}{\epsilon}
\]
Trajectory Generation:

If the state space is very large or even infinite, we cannot solve $X$ numerically (or at least not with standard methods) but we can generate trajectories of $X$ as follows:

Input:
- time horizon $T$
- infinitesimal generator matrix $Q$
- initial distribution $\pi(0)$ (with finitely many nonzero entries)

or a finite representation that can be used to build $Q$ on-the-fly.

Output: trajectory $t \rightarrow x(t,w)$ of $X$.

Remark: Instead of $1, j, k, ...$ we use $ss', ...$ here to denote a state.

1) Initialize current state $s$ according to $\pi(0)$ and set $t=0$. (Choose initial state at random according to probabilities in $\pi(0)$.)

2) Generate a random residence time $\tau$ for state $s$ (i.e., the time that the process remains in state $s$), which follows a negative exponential distribution with parameter $q_s = \sum_{s'} q_{ss'}$

3) Set $t = t + \tau$

4) Choose the next state $s'$ according to the jump probabilities in $s$ (i.e., choose $s'$ at random with probability $q_{ss'}/q_s$)

5) Set $s = s'$

6) If $t < T$ go to step 2), otherwise stop.
Simulate one trajectory

Initialise
while $t < T$

1. generate time until next reaction
2. decide which reaction happens
3. update state
Simulate one trajectory

Initialise
while $t < T$
  1. generate time until next reaction
  2. decide which reaction happens
  3. update state

\[ X = X(0), \ t = 0 \]
Simulate one trajectory

Initialise
while $t < T$

1. generate time until next reaction
2. decide which reaction happens
3. update state

\[
\alpha := \sum_{s=1}^{R} a_s(X(t))
\]
\[
\tau = \text{rand}_\text{exp}(-\alpha)
\]
Simulate one trajectory

Initialise
while $t < T$
1. generate time until next reaction
2. decide which reaction happens
3. update state

$$\alpha := \sum_{s=1}^{R} a_s(X(t))$$

chance of reaction $r$ to happen: $\frac{a_r(X(t))}{\alpha}$
Simulate one trajectory

Initialise

while $t < T$

1. generate time until next reaction
2. decide which reaction happens
3. update state

\[ X(t + \tau) = X(t) + \nu^+_r - \nu^-_r \]
Weight of a trajectory

- Remember: \( \alpha := \sum_{s=1}^{R} a_s(X(t)) \)
- Waiting time: \( \tau = \text{rand}_{\exp}(-\alpha) \) that is \( p(\tau) \sim \alpha e^{-\alpha \tau} \)
- Chance of reaction \( r \) to happen: \( \frac{a_r(X(t))}{\alpha} \)

That gives:

\[
W = \prod_{\text{steps}} \left( \frac{a_r(X(t))}{\alpha} \right) \cdot (\alpha e^{-\alpha \tau}) = \prod_{\text{steps}} a_r(X(t)) e^{-\alpha \tau}
\]
Perturbed weight of a trajectory

- \( J \) is the set of all possible trajectories of length \( T \)
- average:

\[
\mathbb{E}[f(T))] = \frac{\sum_{j \in J} W_j \cdot f_j(T)}{\sum_{j \in J} W_j}
\]
Perturbed weight of a trajectory

- $J$ is the set of all possible trajectories of length $T$

- average:

$$E[f(T)] = \frac{\sum_{j \in J} W_j \cdot f_j(T)}{\sum_{j \in J} W_j}$$

- for changed weights $W'$:

$$E'[f(T)] = \frac{\sum_{j \in J} W'_j \cdot f_j(T)}{\sum_{j \in J} W'_j} = \frac{\sum_{j \in J} (W'_j/W_j)W_j \cdot f_j(T)}{\sum_{j \in J} (W'_j/W_j)W_j}$$
Perturbed weight of a trajectory

- \( J \) is the set of all possible trajectories of length \( T \)
- average:

\[
\mathbb{E}[f(T)] = \frac{\sum_{j \in J} W_j \cdot f_j(T)}{\sum_{j \in J} W_j}
\]

- for changed weights \( W' \):

\[
\mathbb{E}[f(T)'] = \frac{\sum_{j \in J} W'_j \cdot f_j(T)}{\sum_{j \in J} W'_j} = \frac{\sum_{j \in J} (W'_j/W_j)W_j \cdot f_j(T)}{\sum_{j \in J} (W'_j/W_j)W_j}
\]

\[
= \frac{\sum_{j \in J} W_j(W'_j/W_j) \cdot f_j(T)}{\sum_{j \in J} W_j(W'_j/W_j)} = \frac{\mathbb{E}[f(T)W'/W]}{\mathbb{E}[W'/W]}
\]
Estimating derivatives

Differentiating

$$E[f(T)'] = \frac{E[f(T)W'/W]}{E[W'/W]}$$

with respect to $k$ gives:

$$\frac{d}{dk'} E[f(T)'] \bigg|_{k' \to k} = \frac{d}{dk} E[f(T)] = E[f(T) \cdot W_k] - E[f(T)] \cdot E[W_k]$$

with

$$W_k = \frac{d}{dk} \log W$$
Estimating derivatives for our weights

- calculate sensitivity with respect to parameter $k_w \in \{k_1, \ldots, k_R\}$
- remember

$$W_k = \frac{d}{dk} \log W$$

and

$$W = \prod_{\text{steps}} a_r(X(t)) e^{-\alpha \tau} \text{ with } \alpha := \sum_{s=1}^{R} a_s(X(t))$$

- plugging in yields

$$W_{k_w} = \sum_{\text{steps}} \left( \frac{d}{dk_w} \ln a_r(X(t)) - \tau \cdot \frac{d}{dk_w} \alpha \right)$$
Estimating derivatives for our weights

\[ W_{k_w} = \sum_{\text{steps}} \left( \frac{d}{dk_w} \ln a_r(X(t)) - \tau \cdot \frac{d}{dk_w} \alpha \right) \]

\[ = \sum_{\text{steps}} \left( \frac{d}{dk_w} \ln [k_r \cdot h_r(X(t))] - \tau \cdot \frac{d}{dk_w} \sum_{s=1}^{R} [k_s \cdot h_s(X(t))] \right) \]

\[ = \sum_{\text{steps}} \left( \frac{d}{dk_w} \ln [k_r] + \ln [h_r(X(t))] - \tau \cdot h_w(X(t)) \right) \]

\[ = \sum_{\text{steps}} \left( \begin{cases} 1/k_w & \text{if } w = r \\ 0 & \text{if } w \neq r \end{cases} \right) - \tau \cdot h_w(X(t)) \]

\[ = \frac{1}{k_w} N_w(T) - \sum_{\text{steps}} \tau \cdot h_w(X(t)) \]

\[ = \frac{1}{k_w} \left( N_w(T) - \sum_{\text{steps}} \tau \cdot a_w(X(t)) \right) \]
The whole story

$$\frac{d}{dk_w} \mathbb{E}[f(T)] = \mathbb{E}[f(T) \cdot W_{k_w}] - \mathbb{E}[f(T)] \cdot \mathbb{E}[W_{k_w}]$$

with

$$W_{k_w} = \frac{1}{k_w} \left( N_w(T) - \sum_{\text{steps}} \tau \cdot a_w(X(t)) \right)$$

in practise:

- generate a trajectory and calculate $f(T)$, $W_{k_w}$ and $f(T) \cdot W_{k_w}$
- do that many times and average the results
- calculate variance along the way and stop if it is sufficiently small
Simulate one trajectory

Initialise

while $t < T$

1. generate time until next reaction
2. decide which reaction happens
3. update state
4. update weights
Simulate one trajectory

Initialise
while $t < T$
   1. generate time until next reaction
   2. decide which reaction happens
   3. update state
   4. update weights

\[
W_{k_w} = \frac{1}{k_w} \left( N_w(T) - \sum_{\text{steps}} \tau \cdot a_w(X(t)) \right)
\]

watched reaction happened:
\[
W_{\text{new}} = W_{\text{old}} + \frac{1}{k_w} \left( 1 - a_w(X(t)) \cdot \tau \right)
\]

another reaction happened:
\[
W_{\text{new}} = W_{\text{old}} + \frac{1}{k_w} \left( 0 - a_w(X(t)) \cdot \tau \right)
\]
An example: Birth-Death-Process

\[(\ldots) \xrightarrow{k_2} X \xrightarrow{k_1} \emptyset\]

can be solved analytically:

- birth-rate: \(k_2\), death-rate: \(k_1 \cdot X(t)\)
- change of population:

\[
\frac{d}{dt}X(t) = k_2 - k_1 \cdot X(t)
\]

- we start with \(X_0\) molecules
- solution of that ODE:

\[
X(t) = X_0 e^{-k_1 t} + \frac{k_2}{k_1} (1 - e^{-k_1 t})
\]
An example: Birth-Death-Process

\[(\ldots) \xrightarrow{k_2} X \xrightarrow{k_1} \emptyset\]

can be solved analytically:

- we start with \(X_0\) molecules
- solution of that ODE:

\[
X(t) = X_0 e^{-k_1 t} + \frac{k_2}{k_1} (1 - e^{-k_1 t})
\]

two interesting functionals:

- \(X(T)\)
- \(1_{\{a \leq X(T) \leq b\}}\)
$F(X(.)) = X(T)$, Parameter $k_1$
$F(X(\cdot)) = X(T)$, Parameter $k_2$
\[ F(X(.)) = \mathbb{1}_{\{a \leq X(T) \leq b\}} \]
\[ F(X(.)) = 1_{\{1 \leq X(T) \leq 100\}} , \ k_2 \]
\[ F(X(.)) = \mathbb{1}_{\{12 \leq X(T) \leq 13\}} , k_2 \]
References

Questions?

Here are some suggestions:

- How to treat outputs which depend not just on $T$?
- Why is $E[f(T)] = 0$?
- How to differentiate $E[f(T)'] = E[f(T)W'/W]/E[W'/W]$?
- Why not use $p(j) = W_j/\sum_{i\in J} W_i$ as an probability?
- (Why) Is $\sum_{j\in J} W_j < \infty$?