Simulation and optimization in transportation: an overview

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

November 6, 2014
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Transport policies

Complexity

- Transport systems are complex
- Many elements interact
- Presence of uncertainty
Transport policies

Causal effects

- Very important to identify the causal effects
- Failure to do so may generate wrong conclusions
Example: improving safety

Accidents in Kid City

- The mayor of Kid City has commissioned a consulting company
- Objective: assess the effectiveness of safety campaigns
- Before and after analysis
Example: improving safety

Accidents in Kid City
Example: improving safety

Accidents in Kid City:
Example: improving safety

Accidents in Kid City

[Image of a map with various buildings and labeled points]
Example: improving safety

Accidents in Kid City
Example: improving safety

Accidents in Kid City:
Example: improving safety

Conclusions

- The “Drive safely” signs have a significant impact on safety
- The number of accidents has been reduced by 57%, from 21 down to 9.
Example: improving safety

Two major flaws

- Causal effects are not modeled
- Simulation performed with only one draw
Simulation

the act of imitating the behavior of some situation or some process by means of something suitably analogous
Simulation: what it is not in engineering
Simulation

\[ z = h(x, y, u) \]

- Control — \( u \)
- External input — \( y \)

Complex system — state \( x \)

Indicators — \( z \)
Simulation

\[ Z = h(X, Y, U) + \varepsilon_z \]

- **Control** — \( u \)
- **External input** — \( y \)
- **Complex system** — state \( x \)
- **Indicators** — \( z \)

\( \varepsilon_x \), \( \varepsilon_y \), \( \varepsilon_u \), \( \varepsilon_z \)
Simulation

Propagation of uncertainty

\[ Z = h(X, Y, U) + \varepsilon_z \]

- Given the distribution of \( X, Y, U \) and \( \varepsilon_z \)
- What is the distribution of \( Z \)?

Derivation of indicators

- Mean
- Variance
- Modes
- Quantiles
Sampling

- Draw realizations of $X$, $Y$, $U$, $\varepsilon_z$
- Call them $x^r$, $y^r$, $u^r$, $\varepsilon^r_z$
- For each $r$, compute

$$z^r = h(x^r, y^r, u^r) + \varepsilon^r_z$$

- $z^r$ are draws from the random variable $Z$
Empirical distribution function

\[ F_e(x) = \frac{1}{R} \# \{ z^r \leq x \}, \]

For any \( x \in \mathbb{R} \),

\[ \mathbb{E}[F_e(x)] = F(x) \]

and

\[ \text{var}(F_e(x)) = \frac{1}{R} F(x)(1 - F(x)). \]
Statistics

Indicators

- Mean: \( \mathbb{E}[Z] \approx \bar{Z}_R = \frac{1}{R} \sum_{r=1}^{R} z^r \)
- Variance: \( \text{Var}(Z) \approx \frac{1}{R} \sum_{r=1}^{R} (z^r - \bar{Z}_R)^2 \).
- Modes: based on the histogram
- Quantiles: sort and select

Important: there is more than the mean
The mean

The State of the drunk at his AVERAGE position is **ALIVE**

But the AVERAGE State of the drunk is **DEAD**

Savage et al. (2012)
Simulation

The mean

The flaw of averages

Savage et al. (2012)

\[ \mathbb{E}[Z] = \mathbb{E}[h(X, Y, U) + \varepsilon_z] \neq h(\mathbb{E}[X], \mathbb{E}[Y], \mathbb{E}[U]) + \mathbb{E}[\varepsilon_z] \]

... except if \( h \) is linear.
There is more than the mean

Example

- Intersection with capacity 2000 veh/hour
- Traffic light: 30 sec green / 30 sec red
- Constant arrival rate: 2000 veh/hour during 30 minutes
- With 30% probability, capacity at 80%.
- Indicator: Average time spent by travelers
There is more than the mean
Simulation

Pitfalls of simulation

Few number of runs
- Run time is prohibitive
- Tempting to generate partial results rather than no result

Focus on the mean
- The mean is useful, but not sufficient.
- For complex distributions, it may be misleading.
- Intuition from normal distribution (mode = mean, symmetry) do not hold in general.
- Important to investigate the whole distribution.
- Simulation allows to do it easily.
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Optimization

Assumptions

- $U$ is deterministic.
- $S^R(Z)$ is the statistic of $Z$ under interest (mean, quantile, etc.)
- $R$ is the number of draws generated to obtain the statistics
- Distributions of $X$, $Y$ and $\varepsilon_Z$ are known.

Optimization problem

$$\min_u f(u) = S^R(Z) = S^R(h(X, Y, u) + \varepsilon_Z)$$

subject to

$$g(u) = 0.$$
Optimization problem

\[ \min_u f(u) = S^R(Z) = S^R(h(X, Y, u) + \varepsilon_z) \]

subject to
\[ g(u) = 0. \]

Difficulties
- \( R \) must be large, so calculating \( f \) is computationally intensive
- The derivatives of \( f \) are unavailable or very difficult to obtain
Traffic simulation

Parameters calibration

- $X$: state of traffic
- $Y$: observed link flows
- $u$: parameters of the simulator
- $h$: traffic simulator
- $Z$: total squared difference between modeled and observed flows
- $S^R(Z)$: mean squared error
Traffic simulation

Traffic light optimization

- $X$: state of traffic
- $Y$: OD matrices
- $u$: traffic light configuration
- $h$: traffic simulator
- $Z$: total travel time
- $S^R(Z)$: mean of total travel time Osorio and Bierlaire (2013)
- $S^R(Z)$: std. dev. of total travel time Chen et al. (2013)
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Scenario based optimization

Method
- Identify a list of scenarios $u_1, \ldots, u_N$
- Compute $f(u_i)$ for each $i$

Comments
- Solution is feasible and realistic
- Limited computational effort
- No systematic investigation
- Relies only on the creativity of the analyst

"Of course, this is a worst case scenario."
Nonlinear programming

General approach
- $f(u) = S^R(h(X, Y, u) + \varepsilon_z)$ is a nonlinear function of $u$
- In general, it is continuous and differentiable
- As $h$ is a computer program, the derivatives are not available

Methods
- Automatic differentiation  Griewank (2000)
- Derivative-free optimization Conn et al. (2009)
- Direct search  Lewis et al. (2000)
Automatic differentiation

Method

- A software is a sequence of a finite set of elementary operations
- Each of them is easy to differentiate
- Use chain rule to propagate

Derivative-free optimization

Method

- Build a model of the function using interpolation
  - Lagrange polynomials
  - Splines
  - Kriging
- Use a trust region framework to guarantee global convergence

Comments

- Convergence theory
- Numerical issues with interpolation
- Need for a large number of interpolation points
Direct search

Method
- Generate a sequence of simplices using geometrical transformations maintaining the simplex structure

Comments
- Some do not always converge (Nelder-Nead)
- Convergence may be slow
Heuristics

Neighborhood
- Simple modifications of $u$
- Feasible or infeasible

Local search
- Select a better neighbor
- Stop at a local optimum

Meta heuristics
- Escape from local optima
- Simulated annealing
- Variable neighborhood search
- and many others...
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Example of simulation

Machine with 4 states wrt wear

- perfect condition,
- partially damaged,
- seriously damaged,
- completely useless.

Transition:

\[
\begin{pmatrix}
0.95 & 0.04 & 0.01 & 0.0 \\
0.0 & 0.90 & 0.05 & 0.05 \\
0.0 & 0.0 & 0.80 & 0.20 \\
1.0 & 0.0 & 0.0 & 0.0
\end{pmatrix}
\]
Noise reduction: $R = 100$
Noise reduction: \( R = 1000 \)
Noise reduction: $R = 10000$
Noise reduction methods

Adaptive Monte-Carlo

- $R$ varies across iterations
- Small $R$ in early iterations
- $R$ increases as the algorithm converges

Bastin et al. (2006)
Noise reduction methods

Interpolation: true function
Noise reduction methods

Interpolation: simulated function
Noise reduction methods

 Least-square fitting: simulated function
Noise reduction methods

Least square fitting

- Interpolation model + adaptive Monte-Carlo
- Each iterate considered as a sample
- Regression is used instead of interpolation

Comments

- Originally for systems of nonlinear equations
- An update formula à la Broyden can be derived
- Appropriate for large-scale applications (2 millions variables)
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Open box algorithms

What are we simulating?

- $h(\cdot)$ is a detailed description of our system
- We need simulation because it is complicated
- We open the box, and build a simpler representation of the system
Deterministic model

Congestion

- Queuing theory
- Closed form analytical equations
- Simplifying assumptions (e.g. stationarity)

Osorio and Bierlaire (2009)
Metamodel

\[ m(u, x; \alpha, \beta, q) = \alpha T(u, x, q) + \phi(u, \beta) \]

- \( T(\cdot) \) analytical model
- \( \phi(\cdot) \) interpolation model
- \( u \) control (traffic lights)
- \( x \) state variables

Osorio and Bierlaire (2013)
Metamodel

\[ m(u, x; \alpha, \beta, q) = \alpha T(u, x, q) + \phi(u, \beta) \]

- \( T(\cdot) \) analytical model
- \( \phi(\cdot) \) interpolation model
- \( u \) control (traffic lights)
- \( x \) state variables
- engineering
- mathematics
Metamodelling approach

Ongoing research

- **Large scale problems** Osorio and Chong (ta)
- **Fuel consumption** Osorio and Nanduri (ta)
- **Emissions** Osorio and Nanduri (2013)
Large scale problems

Simulated travel time (with 50 draws)  Osorio and Chong (ta)

Initial signal plan

Optimized signal plan
Reliability

Simulated standard deviation (with 50 draws)\textsuperscript{chen et al. (2013)}

Initial signal plan

Optimized signal plan
Outline

1. Simulation
2. Simulation-based optimization
3. Black box algorithms
4. Noise reduction
5. Open box algorithms
6. Conclusions
Summary

Simulation
- Number of draws
- Beyond the mean

Black box algorithms
- Scenarios
- Automatic differentiation
- Derivative-free
- Direct search
- Heuristics
- Noise reduction

Open box algorithms
- Deterministic engineering model
- Metamodel
Conclusion

Everything should be made as simple as possible, but no simpler

Albert Einstein
Bibliography I


Conclusions

Bibliography II


Bibliography III


