Case studies in Optimization

- Monte-Carlo statistical optimization

Candidate Problems for Monte-Carlo

- Problems of the form:

$$\begin{array}{ll} \text{minimize} & \int f_0(x,t)\,dt \\ \text{subject to} & f_i(x) \le 0 \\ & g_j(x) = 0 \end{array} \qquad \begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & \max_v(f_i(x,v)) \le 0 \\ & g_j(x) = 0 \end{array}$$

  are commonly found in real world problems
- Monte-Carlo is good for problems that are too complex to solve directly

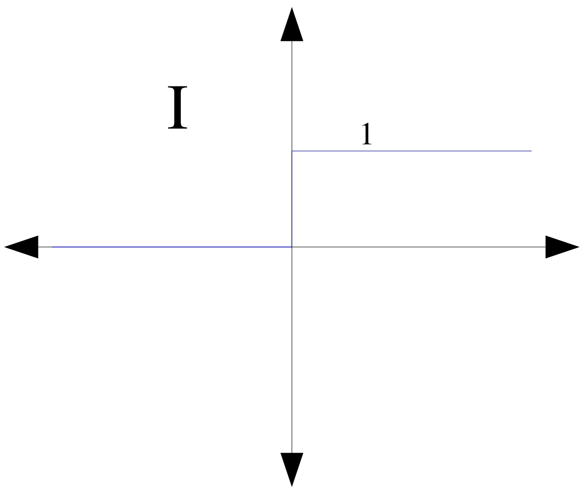
Example Problems

- Example 1: Cost over time

$$\begin{array}{ll} \text{minimize} & \int \text{profit}(x,t)\,dt - \int \text{parts\_cost}(x,t)\,dt \\ \text{subject to} & \text{manfacturable}(x) \geq 0 \end{array}$$

- Example 2: Maximize Yield

$$\begin{array}{ll} \text{maximize} & \int I(\text{specs}(x,v) - s_{\min})\,\mathrm{d}v \\ \text{subject to} & \text{manufacturable}(x) \geq 0 \end{array}$$



Example Problems

- Example 3: Minimize

$$\begin{array}{ll} \text{minimize} & \text{profit}(x) \\ \text{subject to} & \max(\text{risk}(x)) \leq 5 \end{array}$$

- Example 4: Yield Constraint

$$\begin{array}{ll} \text{maximize} & \int \text{profit}(x, v)\,\mathrm{d}v \\ \text{subject to} & \int I(\text{specs}(x, v) - s_{\text{min}})\,\mathrm{d}v \geq 95\,\% \end{array}$$

Monte Carlo & Optimization

Monte-Carlo can be used in a variety of ways

$$\begin{aligned} &\text{minimize} && \int f_0(x,t)\,dt \\ &\text{subject to} && f_i(x) \le 0 \\ & && g_j(x) = 0 \end{aligned}$$

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && \max_v (f_i(x,v)) \le 0 \\ & && g_j(x) = 0 \end{aligned}$$

The function evaluations are

fast	Stochastic Gradient	Scenario
slow	Sample Average Approximation	Scenario

Stochastic Gradient

- Method to solve:

$$\begin{aligned} \text{minimize} \quad & \int f_0(x,t)\,dt \\ \text{subject to} \quad & f_i(x) \le 0 \\ & g_j(x) = 0 \end{aligned}$$

- Best when $\int f_0(x,t)\,dt$ can be evaluated relatively quickly
- Estimate the gradient using Monte-Carlo methods

$$\frac{\partial}{\partial x}\int f_0(x,t)\,dt \approx \sum_{i=1}^{N} \frac{\partial f_0(x,t_i)}{\partial x}$$

  - Use this gradient for optimization

Stochastic Gradient

- Example:

$$\begin{aligned} \text{minimize} \quad & \int \max\{\text{Delay}(x,t) - d_{\min}, 0\}\, dt \\ \text{subject to} \quad & \text{Power}(x) \le P_{\min} \\ & x \in \mathbb{R}^{1000} \end{aligned}$$

- Estimate the gradient using Monte-Carlo methods

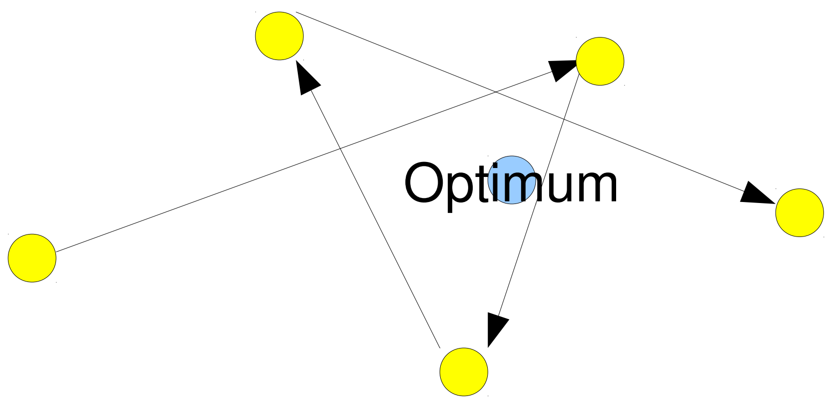
$$\nabla_x \int \max\{\text{Delay}(x,t) - d_{\min}, 0\}\, dt \approx \sum_{i=1}^{N} \nabla_x \max\{\text{Delay}(x,t_i) - d_{\min}, 0\}$$

- Optimize using this approximate gradient (e.g. gradient descent)

$$x^{k+1} = x^k + \alpha \sum_{i=1}^{N} \nabla_x \max\{\text{Delay}(x,t_i) - d_{\min}, 0\}$$

Stochastic Gradient

- How does this work in practice:
  - Expect to get close to the optimum
  - The larger the N, the closer to the optimum



  - Small errors in the gradient computation may prevent convergence
  - Need ~100's of iterations

Sample Average Approximation

Convert the problem:

$$(1)\quad \begin{aligned} \text{minimize} \quad & \int f_0(x,t)\,dt \\ \text{subject to} \quad & f_i(x) \le 0 \\ & g_j(x) = 0 \end{aligned} \quad \xrightarrow{\text{approx}} \quad (2)\quad \begin{aligned} \text{minimize} \quad & \frac{1}{N}\sum_{i=1}^{N} f_0(x,t_i)\,dx \\ \text{subject to} \quad & f_i(x) \le 0 \\ & g_j(x) = 0 \end{aligned}$$

- Fix N samples, and solve the resulting problem
- Useful when $\int f_0(x,t)\,dt$ is difficult to evaluate
- The solutions to (2) are feasible in (1)
- As N gets large, the solutions get better
- This generally works well
  - Current research studies these formulations

Sample Average Approximation

Example

$$\begin{aligned} \text{minimize} \quad & \int \exp(v^T x)\, p(v)\, dv \\ \text{subject to} \quad & \|x - x_0\| \le 0 \\ & x \in \mathbb{R}^{100},\ y \in \mathbb{R}^{100} \end{aligned} \quad \xrightarrow{\text{approx}} \quad \begin{aligned} \text{minimize} \quad & \frac{1}{N} \sum_{i=1}^{N} \exp(v_i^T x) \\ \text{subject to} \quad & \|x - x_0\| \le 0 \\ & x \in \mathbb{R}^{100},\ y \in \mathbb{R}^{100} \end{aligned}$$

- Pick N samples, solve the simplified problem, repeat!
  - Note that $\int \exp(v^T x)\, p(v)\, dv$ is assumed to be too difficult to compute exactly
  - Evaluating $\frac{1}{N} \sum_{i=1}^{N} \exp(v_i^T x)$ is easier
  - Lots of interesting analysis you can do with the solutions

Scenario methods for constraints

Approximate the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & \max_v (f_i(x,v)) \leq 0 \end{array} \xrightarrow{\text{approx}} \begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x, v_n) \leq 0, \;\; \text{n}= 1,\ldots, \text{N} \end{array}$$

- This constraint will be violated with probability ϵ with confidence $1-\beta$, where

$$N \geq \frac{n}{\epsilon \beta} - 1$$

- You can also solve this by adding additional points:
  - Pick the worst v, and add this to the set
  - This will converge

Scenario methods for constraints

Example

$$(2)\quad \begin{array}{ll} \text{minimize} & \text{power}(x) \\ \text{subject to} & \max_v(\text{delay}(x,v)) \leq 0 \end{array}$$

$$\xrightarrow{\text{approx}} \quad (2)\quad \begin{array}{ll} \text{minimize} & \text{power}(x) \\ \text{subject to} & \text{delay}(x, v_n) \leq 0,\ \text{n}= 1,\ldots, \text{N} \end{array}$$

1) Start with $v_{1,} v_{2,} \ldots v_N$

2) Solve (2) get solution x^{N*}

3) Find the worst v

$$v_{\text{worst}} = \text{argmax}\{\text{delay}(x^{N*}, v)\}$$

Add v_{N+1} *to the set*

References

- Shapiro and Philpott, “A tutorial on stochastic programming”
- A. Mutapcic and S. Boyd, “Cutting-set methods for robust convex optimization with pessimizing oracles”
- Calafiore and Campi, “ Randomized Solutions and confidence levels