Data-Driven Distributionally Robust Optimization

Johanna L. Mathieu

Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI, USA

Extreme Thanks to Dr. Bowen Li
Uncertainty in Power System Operations

- Load consumption, renewable energy production $\rightarrow$ need balancing reserves
- Equipment failures $\rightarrow$ need contingency reserves
- How do we choose the right amount of reserves?
  - Experience, historical data
  - Rules of thumb
Coping with Uncertainty

- Stochastic Unit Commitment, Optimal Power Flow (OPF), Economic Dispatch

- One approach: Multistage stochastic optimization over uncertainty scenarios
  - Determine first stage decision variables (e.g., generator dispatch) such that the expected cost over a finite number of uncertainty scenarios affecting second stage decisions (e.g., reserve actions) is minimized subject to constraints repeated for each uncertainty scenario.
  - What happens when what actually happens in real-time doesn’t match any of the uncertainty scenarios?
Another approach: Chance Constrained Optimization

- Used to obtain the lowest cost solution that satisfies constraints at certain (high) probabilities

\[ \mathbb{P}_\xi(f(x, \xi) \leq 0) \geq 1 - \epsilon \]

where \( x \) is the decision variable vector, \( \xi \) is the vector of random variables, and \( \epsilon \) is the user-defined constraint violation probability (usually small e.g., \( \leq 5\% \)).

- For example, in OPF problems physical constraints like line limits should be satisfied for nearly all realizations of wind power production. Vrakopoulou et al. 2013 and Bienstock et al. 2014 incorporate real-time control policies for responding to wind forecast error.
Under certain assumptions you can analytically reformulate chance constraints.

For example, Roald et al. 2013, Bienstock et al. 2014, and Li and Mathieu 2015 assume wind power forecast error is Gaussian and $\epsilon < 50\%$. The resulting constraint is a (convex) second-order cone constraint [Boyd and Vandenberghe 2004].
Wind power forecast error isn’t Gaussian.

Obtaining the actual distribution is impossible – in fact, there is no “actual” distribution.

Empirical studies have shown that assuming the distributions are Gaussian leads to solutions that don’t satisfy the constraint satisfaction threshold. One option is to tune $\epsilon$. 
Alternative Scenario-based Approaches

- Scenario Approach [Campi et al. 2009]
- Probabilistically Robust Approach [Margellos et al. 2014]
- Application to Chance Constrained OPF [Vrakopoulou et al. 2013; Vrakopoulou et al. 2019]
“...in many practical decision situations the raw data can be explained by several strikingly different distributions.” [Delage, Kuhn, Natarajan, Wiesemann 2018]

Distributionally Robust Chance Constraint

$$\inf_{\mathbb{P}_\xi \in \mathcal{D}_\xi} \mathbb{P}_\xi(f(x, \xi) \leq 0) \geq 1 - \epsilon$$

where $\mathcal{D}_\xi$ is an ambiguity set and the chance constraint is satisfied for all distributions within the ambiguity set.

Distributionally Robust OPF [Roald et al. 2015; Zhang et al. 2017; Xie and Ahmed 2017; Summers et al. 2015; Lubin et al. 2016; Li et al. 2017]
Data-Driven Ambiguity Sets

“a family of (possibly infinitely many) probability distributions consistent with the available raw data or prior structural information” [Delage et al. 2018]

Data-based information

- Moments: mean, covariance, higher-order moments
- Distribution structure: support, unimodality, symmetry, log-concavity
- Distance from a reference distribution (e.g., using Wasserstein metric)
Ambiguity set based on known first and second moments

- Ambiguity set:
  
  \[ D_{\xi} := \{ P_{\xi} \in \mathcal{P}^l : \mathbb{E}_{P_{\xi}}[\xi] = \mu, \ \mathbb{E}_{P_{\xi}}[\xi \xi^\top] = \Sigma \} . \]  

- Consider the DR chance constraint
  
  \[ \inf_{P_{\xi} \in D_{\xi}} P_{\xi} (a(x)^\top \xi \leq b(x)) \geq 1 - \epsilon. \]  

- Theorem 2.2 in Wagner derives the exact reformulation of (2) as
  
  \[ \sqrt{\left( \frac{1 - \epsilon}{\epsilon} \right)} a(x)^\top (\Sigma - \mu \mu^\top) a(x) \leq b(x) - a(x)^\top \mu, \]  

  which is a second order cone constraint.
Ambiguity Set Based on Estimated First and Second Moments

- Ambiguity set [Delage and Ye 2010]

\[ \mathcal{D} = \left\{ f(\xi) : \begin{array}{l} \int_{\xi \in \mathbb{R}^K} f(\xi) d\xi = 1 \\ (\mathbb{E}[\xi] - \mu_0)^T \Sigma_0^{-1} (\mathbb{E}[\xi] - \mu_0) \leq \gamma_1 \\ \mathbb{E}[(\xi - \mu_0)(\xi - \mu_0)^T] \preceq \gamma_2 \Sigma_0 \end{array} \right\}, \]

where the three constraints in \( \mathcal{D} \) ensure that (i) the integral of pdf \( f(\xi) \) is one; (ii) the true mean of \( \xi \) lies in a \( \mu_0 \)-centered ellipsoid bounded by \( \gamma_1 \); and (iii) the true covariance matrix lies in a positive semi-definite cone bounded by \( \gamma_2 \Sigma_0 \).

- Jiang and Guan 2015 show how a DR chance constraint with ambiguity set \( \mathcal{D} \) can be reformulated as a semi-definite constraint.
Application to DR OPF with Uncertain Reserves

- For details see Zhang et al. 2017 (on summer school website)
- “Uncertain reserves” from aggregations of residential electric loads providing services like frequency regulation

Table: Cost, Reliability, and CPU Time of A1–A4 for the IEEE 9-Bus System with Congestion ($1 - e_i = 95\%, \forall i$)

<table>
<thead>
<tr>
<th></th>
<th>Gaussian</th>
<th>Scenario</th>
<th>SDP</th>
<th>SOCP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective cost</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>4428.72</td>
<td>4884.12</td>
<td>5030.31</td>
<td>5036.35</td>
</tr>
<tr>
<td>max</td>
<td>4538.47</td>
<td>5036.97</td>
<td>5328.71</td>
<td>5330.85</td>
</tr>
<tr>
<td>min</td>
<td>4345.12</td>
<td>4789.28</td>
<td>4767.10</td>
<td>4767.12</td>
</tr>
<tr>
<td><strong>Reliability (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>78.69</td>
<td>99.44</td>
<td>99.27</td>
<td>99.53</td>
</tr>
<tr>
<td>max</td>
<td>91.94</td>
<td>99.79</td>
<td>99.81</td>
<td>99.85</td>
</tr>
<tr>
<td>min</td>
<td>63.52</td>
<td>99.12</td>
<td>97.93</td>
<td>98.43</td>
</tr>
<tr>
<td><strong>CPU Time (s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>0.05</td>
<td>48.93</td>
<td>6.40</td>
<td>0.47</td>
</tr>
<tr>
<td>max</td>
<td>0.19</td>
<td>49.44</td>
<td>6.77</td>
<td>0.52</td>
</tr>
<tr>
<td>min</td>
<td>0.02</td>
<td>48.22</td>
<td>6.08</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Table: Cost, Reliability, and CPU Time of A1– A4 for the IEEE 39-Bus System with Congestion ($1 - \epsilon_i = 95\%$)

<table>
<thead>
<tr>
<th></th>
<th>Gaussian</th>
<th>Scenario</th>
<th>SDP</th>
<th>SOCP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective cost</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>42350.49</td>
<td>45489.17</td>
<td>44765.94</td>
<td>44787.62</td>
</tr>
<tr>
<td>max</td>
<td>42788.23</td>
<td>45871.69</td>
<td>45676.41</td>
<td>45744.09</td>
</tr>
<tr>
<td>min</td>
<td>41888.16</td>
<td>45187.54</td>
<td>43996.11</td>
<td>44070.00</td>
</tr>
<tr>
<td><strong>Reliability (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>72.38</td>
<td>99.79</td>
<td>93.25</td>
<td>99.10</td>
</tr>
<tr>
<td>max</td>
<td>86.90</td>
<td>99.87</td>
<td>98.23</td>
<td>99.77</td>
</tr>
<tr>
<td>min</td>
<td>60.47</td>
<td>99.70</td>
<td>85.74</td>
<td>96.60</td>
</tr>
<tr>
<td><strong>CPU Time (s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>avg</td>
<td>0.10</td>
<td>1003.45</td>
<td>502.69</td>
<td>0.72</td>
</tr>
<tr>
<td>max</td>
<td>0.12</td>
<td>1040.64</td>
<td>666.80</td>
<td>0.78</td>
</tr>
<tr>
<td>min</td>
<td>0.09</td>
<td>978.80</td>
<td>395.95</td>
<td>0.63</td>
</tr>
</tbody>
</table>

SDP solver having trouble!
Open Problems from Delage et al. 2018

- Choice of performance measure
- Choice of ambiguity set
- Efficiency of solution methods
- Application in multistage settings
- Applications
Choice of performance measure

**Choice of ambiguity set** → Can we obtain less conservative results if we incorporate more structural information? Unimodality, Misspecified modes, Log-concavity

**Efficiency of solution methods** → When we incorporate structural information can we exactly reformulate the problem or do we have to resort to approximations? What are the computational requirements of the reformulations/approximations?

Application in multistage settings

**Applications** → Do the performance/computation trade-offs make sense for OPF problems?
Exercise Set-Up: dc Optimal Power Flow with Wind Power Uncertainty [Vrakopoulou et al. 2013]

Decisions: generation $P_G$, reserve capacities $R^{up}$ and $R^{dn}$, participation factors $d_G$

Random variable: wind forecast error $\tilde{W}$

$$\min P_G^T[C_1]P_G + C_2^T P_G + C_R^T (R^{up} + R^{dn})$$

s.t. $-P_I \leq AP_{inj} \leq P_I$

$$P_{inj} = C_G(P_G + R) + C_W(P_W^f + \tilde{W}) - C_L P_L$$

$$R = -d_G \sum_i \tilde{W}_i$$

$$P_G \leq P_G + R \leq \bar{P}_G$$

$$-R^{dn} \leq R \leq R^{up}$$

$$1_{1 \times N_G} d_G = 1$$

$$1_{1 \times N_B} (C_G P_G + C_W P_W^f - C_L P_L) = 0$$

$$P_G \geq 0_{N_G \times 1}, \; d_G \geq 0_{N_G \times 1}$$

$$R^{up} \geq 0_{N_G \times 1}, \; R^{dn} \geq 0_{N_G \times 1}$$
Notation

- $C_1, C_2, C_R$ are cost parameters
- $P_{\text{inj}}$ is a vector of bus power injections
- $P_W^f$ is a vector of wind power forecasts
- $C_G, C_W, C_L$ are matrices that map generators, wind power plants, and loads to their bus
- $A$ is a matrix derived from network reactances which maps bus power injections to power flows
- $P_l$ is a vector of line flow limits
- $R$ is a vector of real-time reserve actions
- $P_G$ and $\bar{P}_G$ are generator limits
- $\mathbf{1}$ is a vector of ones; the dimension is also given
- $\mathbf{0}$ is a vector of zeros; the dimension is also given
- $N_G, N_B$ are the number of generators, buses
Exercise: Distributionally Robust dc Optimal Power Flow with Wind Power Uncertainty

- What you will do: Reformulate the dc OPF problem as a distributionally robust OPF problem assuming known first and second moments. Solve with MATLAB. Explore what happens when you change the parameters.

- What you are given: Three MATLAB files ... 1) IEEE 9 bus data from MATPOWER (case9.m), 2) a function to read in the data (Caseln.m), and code implementing most of the optimization problem with comments indicating what is left for you to fill in (DR_moment_incomplete.m). Put all files in the same folder along with CVX (or add to path). Modify ONLY DR_moment_incomplete.m.
Exercise: Distributionally Robust dc Optimal Power Flow with Wind Power Uncertainty

1. Form groups of 2-3 students. Familiarize yourself with the dc OPF problem on slide 16. Ask questions if something isn’t clear.

2. Separate the constraints into the deterministic constraints and stochastic constraints. Write the stochastic constraints as chance constraints. Then write the stochastic constraints as distributionally robust constraints.

3. Reformulate the DR constraints as SOC constraints by assuming known first and second moments.

4. Familiarize yourself with the code in DR_moment_incomplete.m.

5. Complete the code and report the optimal cost and decision variables.

6. If you have time, explore what happens when you change the wind forecasts and error statistics, violation probability $\epsilon$, and test system parameters.