Ambiguous Risk Constraints with Moment and Structural Information

Ruiwei Jiang

University of Michigan

BIRS Workshop on Distributionally Robust Optimization

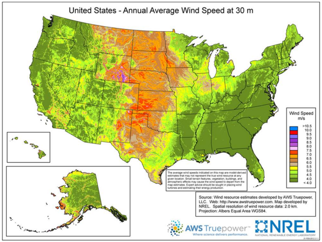
Joint work with Yuanyuan Guo, Bowen Li, and Johanna L. Mathieu, supported by the NSF (CMMI-1662774).

Outline

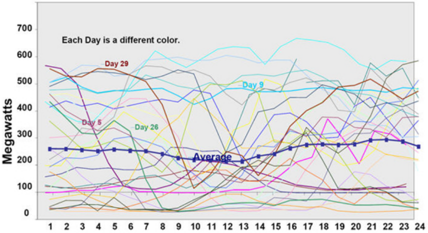
Outline

Application: Integrating Renewable Energy

- Example: wind power.
- Positive: low generation cost and environmentally friendly.
- Negative: intermittent nature.
  - 20% day-ahead prediction MAE for a single wind farm. [NREL, 2015]



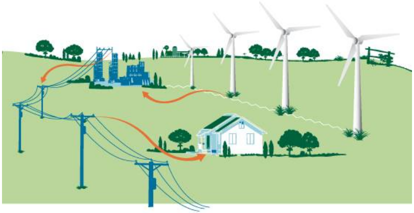
(a) Wind power resource



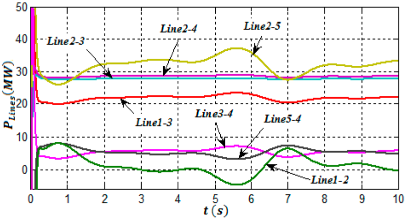
(b) Wind power fluctuation

Application: Integrating Renewable Energy

- Random wind power $\Longrightarrow$ random transmission line flow.
- Risk of line overflow.



(a) Wind to Grid

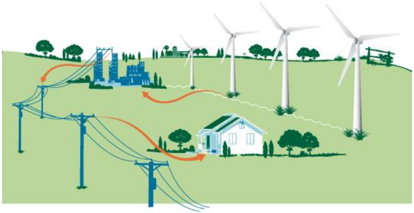


(b) Random Line Flow

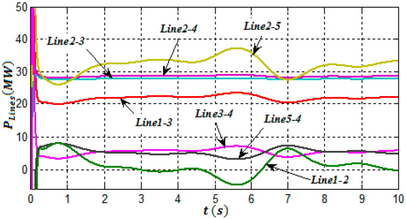
Application: Integrating Renewable Energy

- Random wind power $\Longrightarrow$ random transmission line flow.
- Risk of line overflow.
- DC approximation: line flow $=$ Affine(x, ξ)
  - x: generation scheduling decisions.
  - ξ: wind prediction errors.
- How to control the risk of overflow, i.e.,

$$\text{Affine}(x, \xi) \; > \; \text{Capacity?}$$



(a) Wind to Grid



(b) Random Line Flow

Constraints under Uncertainty

$$a(x)^\top \xi \le b(x)$$

- x: decision variables.
- $a(x)$, $b(x)$: affine functions of x.
- ξ: random vector.

- Constraints under uncertainty in other applications, e.g.,

Constraints under Uncertainty

$$a(x)^\top \xi \leq b(x)$$

- x: decision variables.
- $a(x)$, $b(x)$: affine functions of x.
- ξ: random vector.

- Constraints under uncertainty in other applications, e.g.,
  - (Inventory control) End inventory ≥ 0.
  - (Appointment scheduling) Overtime $\leq T$.

Risk Constraints

$$\mathbb{P}\left\{a(x)^\top \xi \leq b(x)\right\} \ \geq \ 1 - \epsilon$$

- x: decision variables.
- $a(x)$, $b(x)$: affine functions of x.
- ξ: random vector.

- Chance constraints:

Risk Constraints

$$\mathbb{P}\left\{a(x)^\top \xi \le b(x)\right\} \;\ge\; 1-\epsilon$$

- x: decision variables.
- $a(x)$, $b(x)$: affine functions of x.
- ξ: random vector.

- Chance constraints:
  - Probability of violation $\le \epsilon$ (e.g., $\epsilon = 0.05$).
  - Dating back to [Charnes et al., 1958], [Miller and Wagner, 1965].

Risk Constraints

$$\mathbb{P}\left\{a(x)^\top \xi \leq b(x)\right\} \;\geq\; 1-\epsilon$$

- x: decision variables.
- $a(x)$, $b(x)$: affine functions of x.
- ξ: random vector.

- Chance constraints:
  - Probability of violation $\leq \epsilon$ (e.g., $\epsilon = 0.05$).
  - Dating back to [Charnes et al., 1958], [Miller and Wagner, 1965].
    - ★ Production Planning: [Gade and Küçükyavuz, 2013].
    - ★ Chemical Processing: [Henrion and Möller, 2003].
    - ★ Power System Operations: [Ozturk et al., 2004].
    - ★ $\mathbb{P}\{$Line Overflow$\} \approx$ Fraction of Time of Line Overflow. [Bienstock et al., 2014]

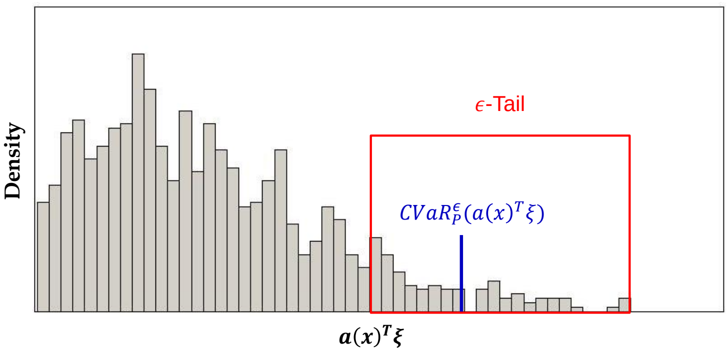
Risk Constraints

- Violation magnitude?
  - $a(x)^\top \xi - b(x)$, given that $a(x)^\top \xi > b(x)$.
  - Chance constraints offer no guarantees on the magnitude.
  - Conditional Value-at-Risk (CVaR) is a natural alternative.

Risk Constraints

- Violation magnitude?
  - $a(x)^\top \xi - b(x)$, given that $a(x)^\top \xi > b(x)$.
  - Chance constraints offer no guarantees on the magnitude.
  - Conditional Value-at-Risk (CVaR) is a natural alternative.
    - Seminal work: [Artzner et al., 1999], [Rockafellar and Uryasev, 2000], [Nemirovski and Shapiro, 2006].
    - Conditional expectation on the upper-ϵ tail.

An Illustration of the CVaR



- CVaR = upper ϵ-tail conditional expectation.

An Application on Integrating Renewable Energy

- How to control the risk of

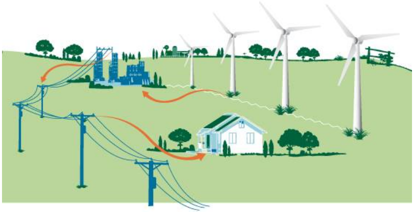
$$\text{Random Transmissoin line Flow} > \text{Capacity?}$$

- Chance constraint:

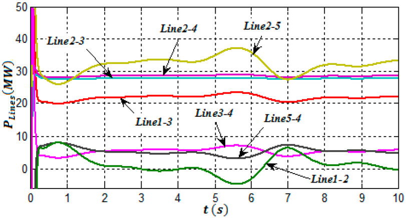
$$\mathbb{P}\left\{\text{Affine}(x,\xi) \leq \text{Capacity}\right\} \;\geq\; 1-\epsilon.$$

- CVaR constraint:

$$\text{CVaR}_{\mathbb{P}}^{\epsilon}\left(\text{Affine}(x,\xi)\right) \;\leq\; \text{Capacity}.$$



(a) Wind to Grid



(b) Random Line Flow

Challenges on Modeling: Imperfect Distributional Info

- $\mathbb{P}$ may not be accurately estimated.
  - Multiple plausible choices.

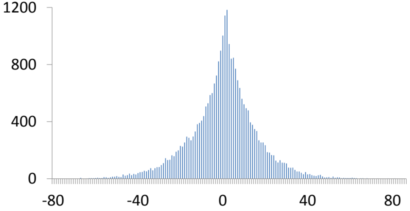


Figure: Prediction Error Histogram

Challenges on Modeling: Imperfect Distributional Info

- $\mathbb{P}$ may not be accurately estimated.
  - Multiple plausible choices.
- Example: wind prediction errors.
  - Normal. [Doherty and O'Malley, 2005]
  - Weibull. [Dietrich et al., 2009]
  - Cauchy. [Hodge and Milligan, 2011]
  - Hyperbolic. [Hodge et al., 2012]

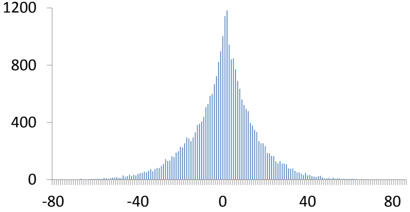


Figure: Prediction Error Histogram

Ambiguous Risk Constraints

- Addressing the distributional ambiguity.
  - A reviving area:
  - Origin (TBMK): [Scarf, 1958]
  - 2000–2010:

    [Shapiro and Kleywegt, 2002], [Nemirovski and Shapiro, 2006], [Goh and Sim, 2010], [Bertsimas et al., 2010], [Delage and Ye, 2010], and more.
  - 2010+:

    [Xu and Mannor, 2012], [Ahmed and Papageorgiou, 2013], [Zymler et al., 2013], [Toriello et al., 2014], [Wiesemann et al., 2014], [Zhao and Guan, 2014], [Yu and Xu, 2015], [Esfahani and Kuhn, 2015], [Yang and Xu, 2016], [Gao and Kleywegt, 2016], [Shapiro, 2016], [Xie and Ahmed, 2016a], [Shapiro, 2017], and many more.

Ambiguous Risk Constraints

- Addressing the distributional ambiguity.
  - A reviving area:
  - Origin (TBMK): [Scarf, 1958]
  - 2000–2010:

    [Shapiro and Kleywegt, 2002], [Nemirovski and Shapiro, 2006], [Goh and Sim, 2010], [Bertsimas et al., 2010], [Delage and Ye, 2010], and more.
  - 2010+:

    [Xu and Mannor, 2012], [Ahmed and Papageorgiou, 2013], [Zymler et al., 2013], [Toriello et al., 2014], [Wiesemann et al., 2014], [Zhao and Guan, 2014], [Yu and Xu, 2015], [Esfahani and Kuhn, 2015], [Yang and Xu, 2016], [Gao and Kleywegt, 2016], [Shapiro, 2016], [Xie and Ahmed, 2016a], [Shapiro, 2017], and many more.
  - A family of probability distributions.
  - Moment-based ambiguity set:

$$\mathcal{D}(\mu, \Sigma) \ = \ \{\mathbb{P} : \ \mathbb{E}_{\mathbb{P}}[\xi] = \mu, \ \mathbb{E}_{\mathbb{P}}[\xi\xi^\top] = \Sigma\}.$$

This Talk

- One step further: moment + structural information.
  - Log-concavity.
  - Tail dominance.

This Talk

- One step further: moment + structural information.
  - Log-concavity.
  - Tail dominance.
- New ambiguity set:

$$\mathcal{D}_{\mathsf{S}}(\mu, \Sigma) \; = \; \mathcal{D}_{\mathsf{S}} \; \cap \; \mathcal{D}(\mu, \Sigma).$$

- Domain knowledge + data-driven.

This Talk

- One step further: moment + structural information.
  - Log-concavity.
  - Tail dominance.
- New ambiguity set:

$$\mathcal{D}_{\mathsf{S}}(\mu, \Sigma) \; = \; \mathcal{D}_{\mathsf{S}} \; \cap \; \mathcal{D}(\mu, \Sigma).$$

- Domain knowledge + data-driven.
- Ambiguous chance constraints (ACC):

$$\inf_{\mathbb{P} \in \mathcal{D}_{\mathsf{S}}(\mu, \Sigma)} \mathbb{P}\left\{a(x)^\top \xi \leq b(x)\right\} \; \geq \; 1 - \epsilon.$$

- Ambiguous CVaR constraints (AVC):

$$\sup_{\mathbb{P} \in \mathcal{D}_{\mathsf{S}}(\mu, \Sigma)} \mathrm{CVaR}_{\mathbb{P}}^{\epsilon}\left(a(x)^\top \xi\right) \; \leq \; b(x).$$

Outline

Motivation

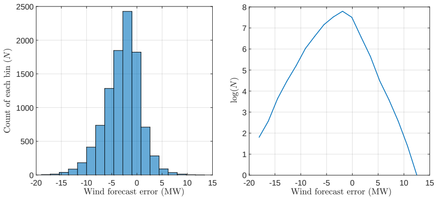
- Example: Wind prediction errors.
  - Normal. [Doherty and O'Malley, 2005]
  - Weibull. [Dietrich et al., 2009]
  - Cauchy. [Hodge and Milligan, 2011]
  - Hyperbolic. [Hodge et al., 2012]

Motivation

- Example: Wind prediction errors.
  - Normal. [Doherty and O'Malley, 2005] (log-concave)
  - Weibull. [Dietrich et al., 2009] (log-concave if shape parameter ≥ 1)
  - Cauchy. [Hodge and Milligan, 2011] (NOT log-concave)
  - Hyperbolic. [Hodge et al., 2012] (log-concave)
- Most of the above candidates are log-concave.

Motivation

- Example: Wind prediction errors.
  - Normal. [Doherty and O'Malley, 2005] (log-concave)
  - Weibull. [Dietrich et al., 2009] (log-concave if shape parameter ≥ 1)
  - Cauchy. [Hodge and Milligan, 2011] (NOT log-concave)
  - Hyperbolic. [Hodge et al., 2012] (log-concave)
- Most of the above candidates are log-concave.



Ambiguity Set with Log-Concave Information

- Log-concave: the log-density function is concave.

Ambiguity Set with Log-Concave Information

- Log-concave: the log-density function is concave.
  - In this example,

$$\mathcal{D}_{\mathsf{S}}(\mu, \Sigma) := \Big\{ \mathbb{P} : \ \mathbb{E}_{\mathbb{P}}[\xi] = \mu, \\ ||\Sigma^{-1/2}(\xi - \mu)||_2 \le r \ \text{ almost surely}, \\ \mathbb{P} \text{ is log-concave} \Big\}.$$

  - Mean, support, and log-concave structural information.
  - We consider log-concave density.

Log-Concavity Density $\Rightarrow$ CDF Log-Concavity

Related Work

- Classical results on the convexity of (non-ambiguous) chance constraints

$$\mathbb{P}\{c^\top x + d \le 0\} \ \ge \ 1 - \epsilon.$$

- Uncertainty quantification of remaining lifetime in reliability literature:

$$\mathbb{P}\{X > t\}.$$

Related Work

- Classical results on the convexity of (non-ambiguous) chance constraints
$$\mathbb{P}\{c^\top x + d \le 0\} \ \ge \ 1 - \epsilon.$$
  - $(c^\top, d)$ Gaussian $\Rightarrow$ SOC representation. [van de Panne and Popp, 1963].
  - c deterministic, d log-concave $\Rightarrow$ convex. [Prékopa, 1995].
  - $(c^\top, d)$ log-concave and symmetric $\Rightarrow$ convex. [Lagoa et al., 2001].
- Uncertainty quantification of remaining lifetime in reliability literature:
$$\mathbb{P}\{X > t\}.$$

Related Work

- Classical results on the convexity of (non-ambiguous) chance constraints

$$\mathbb{P}\{c^\top x + d \le 0\} \;\ge\; 1 - \epsilon.$$

  - $(c^\top, d)$ Gaussian $\Rightarrow$ SOC representation. [van de Panne and Popp, 1963].
  - c deterministic, d log-concave $\Rightarrow$ convex. [Prékopa, 1995].
  - $(c^\top, d)$ log-concave and symmetric $\Rightarrow$ convex. [Lagoa et al., 2001].

- Uncertainty quantification of remaining lifetime in reliability literature:

$$\mathbb{P}\{X > t\}.$$

  - Sharp upper bound if the CDF of X is log-concave and $\mathbb{E}_{\mathbb{P}}[X^r]$ is known. [Sengupta and Nanda, 1998]

Related Work

- Classical results on the convexity of (non-ambiguous) chance constraints
$$\mathbb{P}\{c^\top x + d \leq 0\} \;\geq\; 1 - \epsilon.$$
  - $(c^\top, d)$ Gaussian $\Rightarrow$ SOC representation. [van de Panne and Popp, 1963].
  - c deterministic, d log-concave $\Rightarrow$ convex. [Prékopa, 1995].
  - $(c^\top, d)$ log-concave and symmetric $\Rightarrow$ convex. [Lagoa et al., 2001].
- Uncertainty quantification of remaining lifetime in reliability literature:
$$\mathbb{P}\{X > t\}.$$
  - Sharp upper bound if the CDF of X is log-concave and $\mathbb{E}_{\mathbb{P}}[X^r]$ is known. [Sengupta and Nanda, 1998]
- Our focus: DRO among all log-concave densities.
- Results: SOC conservative approximations of ACC and AVC.

Main Results – ACC Approximation I

Theorem: SOC Conservative Approximation for ACC

Under moment and log-concavity information, and if $\epsilon < 1/4$, then ACC

$$\inf_{\mathbb{P} \in \mathcal{D}_S(\mu, \Sigma)} \mathbb{P}\left\{a(x)^\top \xi \leq b(x)\right\} \ \geq \ 1 - \epsilon$$

is implied by the SOC constraint:

$$\mu^\top a(x) + r\left[1 - \frac{2\log(1-\epsilon)}{d^*}\right]\left\|\Sigma^{1/2} a(x)\right\|_2 \ \leq \ b(x),$$

where d^* represents the unique root of function $e^d - d/2 - 1$ on the interval $(-\infty, 0)$.

- Obtained by relaxing the (PDF) log-concavity to the CDF log-concavity of $\mathbb{P}$.

Main Results – ACC Approximation II

Theorem: SOC Conservative Approximation for ACC

Under moment and log-concavity information, and if $\epsilon < 1/4$, then ACC

$$\inf_{\mathbb{P}\in\mathcal{D}_S(\mu,\Sigma)} \mathbb{P}\left\{a(x)^\top \xi \le b(x)\right\} \;\ge\; 1-\epsilon$$

is implied by the SOC constraint:

$$\mu^\top a(x) + \frac{(1-\epsilon)r}{1+\epsilon}\left\|\Sigma^{1/2}a(x)\right\|_2 \;\le\; b(x).$$

- Obtained by relaxing the (PDF) log-concavity to the unimodality of $\mathbb{P}$.
- Existing results on ACC with moment and unimodality information.[1]
- Tighter approximation than the CDF-log-concave one.

[1] Li, B., Jiang, R., Mathieu, J. L., "Ambiguous Risk Constraints with Moment and Unimodality Information," Mathematical Programming, 2018.

Main Results – ACC Approximation III

Theorem: SOC Relaxing Approximation for ACC

Under moment and log-concavity information, ACC

$$\inf_{\mathbb{P}\in\mathcal{D}_{\mathsf{S}}(\mu,\Sigma)} \mathbb{P}\left\{a(x)^\top \xi \leq b(x)\right\} \geq 1-\epsilon$$

implies the SOC constraint:

$$\mu^\top a(x) + r(1-2\epsilon)\left\|\Sigma^{1/2}a(x)\right\|_2 \leq b(x).$$

- Obtained by assuming that $\mathbb{P}$ is uniform.

Main Results – AVC Reformulation

Theorem: SOC Reformulation for AVC

Under moment and log-concavity information, AVC

$$\sup_{\mathbb{P} \in \mathcal{D}_S(\mu, \Sigma)} \text{CVaR}\left(a(x)^\top \xi\right) \leq b(x)$$

is equivalent to the SOC constraint:

$$\mu^\top a(x) + r(1-\epsilon)\left\|\Sigma^{1/2} a(x)\right\|_2 \leq b(x).$$

- Relaxing the (PDF) log-concavity to the unimodality of $\mathbb{P}$.
- Conservative approximation.
- But the worst-case CVaR is attained when $\mathbb{P}$ is uniform!

Extension – Incorporating Covariance I

- Incorporating the covariance information:

$$\mathcal{D}_{\mathsf{S}}(\mu, \Sigma) := \Big\{ \mathbb{P} : \ \begin{aligned}[t] & \mathbb{E}_{\mathbb{P}}[\xi] = \mu, \\ & \mathbb{E}_{\mathbb{P}}[\xi\xi^\top] = \Sigma, \\ & \mathbb{P} \text{ is log-concave} \Big\}. \end{aligned}$$

- Mean, covariance, and log-concave structural information.

Extension – Incorporating Covariance II

Theorem: SOC Conservative Approximation for ACC

Under moment and log-concavity information, ACC

$$\inf_{\mathbb{P} \in \mathcal{D}_{\mathsf{S}}(\mu,\Sigma)} \mathbb{P}\left\{ a(x)^\top \xi \le b(x) \right\} \;\ge\; 1 - \epsilon$$

is implied by the SOC constraint:

$$\mu^\top a(x) + \tau(\epsilon) \left\| \left(\Sigma - \mu\mu^\top \right)^{1/2} a(x) \right\|_2 \;\le\; b(x),$$

where $\tau(\epsilon) = \max\left\{ \sqrt{\frac{3-3\epsilon}{1+3\epsilon}},\ \sqrt{\frac{4}{9\epsilon} - 1} \right\}$.

- Obtained by relaxing the (PDF) log-concavity to the unimodality of $\mathbb{P}$.
- Existing results on ACC with moment and unimodality information.

Extension – Incorporating Covariance III

- Actually, already known as the one-sided VysochanskijPetunin inequality.
- See also [Roald et al. 2015].

Extension – Incorporating Covariance III

- Actually, already known as the one-sided VysochanskijPetunin inequality.
- See also [Roald et al. 2015].
- An independent proof (in English)...

Теорема. Нехай $\bar{U}_d$ — клас одновершинно розподілених випадкових величин ξ з скінченними математичними сподіваннями $\mathbf{M}\xi$ і дисперсіями $\mathbf{D}\xi = d$, де d — деяке фіксоване число. Тоді для всіх $\varepsilon > 0$ виконується рівність

$$\max_{\xi\in U_d} \mathbf{P}(\xi \geqslant \mathbf{M}\xi + \varepsilon) = \begin{cases} (3d - \varepsilon^2)[3(d + \varepsilon^2)]^{-1} \text{ при } \varepsilon^2 \leqslant 5d\cdot 3^{-1}, \\ 4d[9(d + \varepsilon^2)]^{-1} \text{ при } \varepsilon^2 \geqslant 5d\cdot 3^{-1}. \end{cases} \tag{2}$$

Доведення 1. Зафіксуємо розподіл довільної випадкової величини $\xi_0 \in U_d$, яка має нульове математичне сподівання. Тоді, з огляду на [2] і [6, с. 64], при всіх $m \in \mathbf{R}$ для випадкової величини

$$\xi = \xi_0 + m \tag{3}$$

справедлива нерівність

$$\mathbf{P}(\xi \geqslant \mathbf{M}\xi + \varepsilon) \leqslant \max\left\{\frac{4(m^2 + d) - (m + \varepsilon)^2}{3(m + \varepsilon)^2}, \frac{4(m^2 + d)}{9(m + \varepsilon)^2}\right\} = f(m), \tag{4}$$

у лівій частині якої ймовірність $\mathbf{P}(\xi \geqslant \mathbf{M}\xi + \varepsilon) = \mathbf{P}(\xi_0 + m \geqslant m + \varepsilon) = \mathbf{P}(\xi_0 \geqslant \varepsilon)$ (див. [3]). Тому $\forall m \in \mathbf{R}$: $\mathbf{P}(\xi_0 \geqslant \varepsilon) \leqslant f(m)$, де ймовірність не залежить від m. Отже,

$$\mathbf{P}(\xi_0 \geqslant \varepsilon) \leqslant f(d\varepsilon^{-1}). \tag{5}$$

Обчислимо праву частину нерівності (5). Із рівності (4) маємо

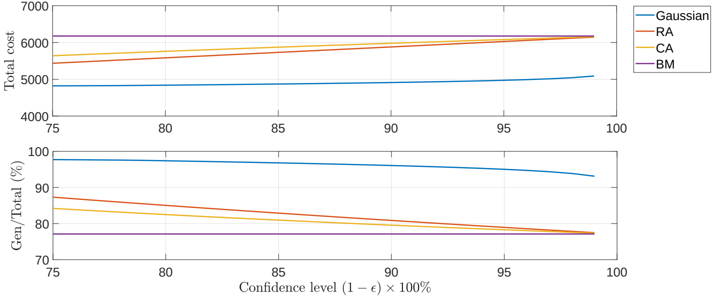
$$f(d\varepsilon^{-1}) = \max\{[4g(d\varepsilon^{-1}) - 1]\cdot 3^{-1}, \quad 4g(d\varepsilon^{-1})\cdot 9^{-1}\}, \tag{6}$$

Figure: (Screenshot) D. Vysochanskij and Y. Petunin, "Improvement of the unilateral 3σ-rule for unimodal distributions," Dokl. Akad. Nauk. Ukr. SSR, Ser. A, vol. 1, pp. 68, 1985.

An Application to Risk-Constrained OPF

- Optimal power flow with wind power.
- IEEE 9-bus system.
- Electricity loads increased by 50%.
- 2 wind farms at buses 2 and 8, respectively.
- Forecasted wind power = 66.8MW.
- Mean and support of forecast errors from historical data.
- ACC on transmission line capacity, upward/downward reserves, and lower/upper bounds of generation amounts.

Optimal Value as ϵ varies



- Gaussian: $\mathbb{P}$ assumed to be Gaussian.
- RA: relaxing approximation.
- CA: conservative approximation.
- BM: benchmark with mean and support information but without log-concavity.

Out-of-Sample Reliability as ϵ varies

Table: Out-of-Sample Reliability (%) with Data Size 500

$1-\epsilon$		Gaussian	RA	CA	BM
	min	81.2	93.1	93.3	95.5
95%	avg	82.3	94.7	94.9	96.7
	max	84.2	96.1	96.1	97.4
	min	50.2	70.5	79.7	95.5
75%	avg	52.3	72.2	81.0	96.7
	max	54.1	74.0	83.1	97.4

- Gaussian not very reliable.
- RA less conservative than BM.
- Further reducing conservatism: incorporating the covariance info.

Outline

Motivation

- One may have more distributional info than the first 2 moments:
  - Directly incorporated into $\mathcal{D}_{\mathsf{S}}$: challenging.
  - Implies Markov-like inequalities.

Motivation

- One may have more distributional info than the first 2 moments:
  - Directly incorporated into $\mathcal{D}_S$: challenging.
  - Implies Markov-like inequalities.
- Higher moments.

$$\mathbb{P}\{|X-\mu| \geq t\} \ \leq \ \frac{\mathbb{E}_{\mathbb{P}}[|X-\mu|^k]}{t^k}, \quad \forall t > 0.$$

- Sub-Gaussian.

$$\mathbb{P}\{|X-\mu| \geq t\} \ \leq \ c\mathbb{P}\{|\mathcal{N}(\mu,\tau^2)-\mu| \geq t\}, \quad \forall t \geq 0.$$

- Sub-exponential.

$$\mathbb{P}\{|X-\mu| \geq t\} \ \leq \ c_1 e^{-c_2 t}, \quad \forall t > 0.$$

- Random vector: Hanson-Wright inequalities. [Hanson and Wright, 1971]

Motivation

- In this example,

$$\mathcal{D}_{\mathsf{S}}(\mu, \Sigma) := \Big\{ \mathbb{P} : \ \mathbb{E}_{\mathbb{P}}[\xi] = \mu, \\ \mathbb{P}\big\{ ||\Sigma^{-1/2}(\xi - \mu)|| > r \big\} \leq \epsilon(r), \ \forall r \in [r_{\mathsf{L}}, r_{\mathsf{U}}], \\ ||\Sigma^{-1/2}(\xi - \mu)|| \leq \bar{r} \ \text{ almost surely} \Big\}.$$

- Mean, support, and dominance information.
- Examples of $\epsilon(r)$:
  - Higher moments: $\mathbb{E}_{\mathbb{P}}[|X - \mu|^k]/r^k$.
  - Sub-Gaussian: $c\mathbb{P}\{|\mathcal{N}(\mu, \tau^2) - \mu| > r\}$.
  - Sub-exponential: $c_1 e^{-c_2 r}$.
- Extensions:
  - Incorporate covariance matrix.
  - Replace $||\Sigma^{-1/2}(\xi - \mu)||$ with a general distance $d(\xi, \mu)$.

Main Results – Worst-Case Expectation

Theorem: Upper Bound

For a general function $f(x, \xi)$,

$$\sup_{\mathbb{P} \in \mathcal{D}_{\mathsf{S}}(\mu, \Sigma)} \mathbb{E}_{\mathbb{P}}[f(x, \xi)] \ \leq \ \min_{p} \ \mathbb{E}_{\mathbb{Q}}[H(p, \zeta)],$$

where

$$H(p, \zeta) := \max_{\xi : ||\Sigma^{-1/2}(\xi - \mu)|| \leq \zeta} \left\{ f(x, \xi) - p^{\top}(\xi - \mu) \right\}$$

and ζ represents a random variable and $\mathbb{Q}$ represents its CDF:

$$\mathbb{Q}\{\zeta \leq x\} \ = \begin{cases} 0, & \text{if } x < r_{\mathsf{L}}, \\ 1 - \epsilon(x), & \text{if } r_{\mathsf{L}} \leq x \leq r_{\mathsf{U}}, \\ 1 - \epsilon(r_{\mathsf{U}}), & \text{if } r_{\mathsf{U}} < x < \bar{r}, \\ 1, & \text{if } x \geq \bar{r}. \end{cases}$$

Main Results – Worst-Case Expectation

Theorem: Tightness

If $f(x,\xi)$ can be written as the maximum of functions concave in ξ, i.e., there exist $f_i(x,\xi)$, $\forall i \in [I]$, concave in ξ such that

$$f(x,\xi) \;=\; \max_{i\in[I]}\left\{f_k(x,\xi)\right\},$$

then

$$\sup_{\mathbb{P}\in\mathcal{D}_S(\mu,\Sigma)} \mathbb{E}_{\mathbb{P}}[f(x,\xi)] \;=\; \min_{p}\; \mathbb{E}_{\mathbb{Q}}[H(p,\zeta)].$$

- Most relevant case in applications: $f(x,\xi) = \max_{i\in[I]}\left\{a_i(x)^\top\xi - b_i(x)\right\}$.
  - Newsvendor, AVC, two-stage DR stochastic linear programming.
- Uncertainty quantification $\sup_{\mathbb{P}\in\mathcal{D}_S} \mathbb{P}\{\exists i \in [I] :\; a_i(x)^\top\xi > b_i(x)\}$:

$$f(x,\xi) = \max_{i\in[I]}\left\{\chi_{[a_i(x)^\top\xi > b_i(x)]}(\xi)\right\}.$$

Main Results – Worst-Case Expectation

Theorem: Most Relevant Case

If $f(x,\xi) := \max_{i\in[I]}\left\{a_i(x)^\top\xi - b_i(x)\right\}$, then

$$\sup_{\mathbb{P}\in\mathcal{D}_{\mathsf{S}}(\mu,\Sigma)} \mathbb{E}_{\mathbb{P}}[f(x,\xi)] \;=\; \min_{p}\; \mathbb{E}_{\mathbb{Q}}\left[\max_{i\in[I]}\left\{||\Sigma^{1/2}(a_i(x)-p)||_*\zeta + \mu^\top a_i(x) - b_i(x)\right\}\right].$$

- Conclusion valid for $\mathbb{Q}$ being continuous or discrete.
- Worst-case distributions available.
- Reformulation jointly convex in (x,p).

Main Results – Worst-Case Expectation

Theorem: Special Case

If $f(x, \xi)$ is concave in ξ, then

$$\sup_{\mathbb{P} \in \mathcal{D}_S(\mu, \Sigma)} \mathbb{E}_{\mathbb{P}}[f(x, \xi)] \; = \; f(x, \mu).$$

- Worst-case distribution is supported at μ.

Main Results – Worst-Case Expectation

Theorem: What if $f(x,\xi)$ is convex in ξ?

$$\mathbb{E}_{\mathbb{P}}[(\xi-\mu)(\xi-\mu)^\top] \preceq \mathbb{E}_{\mathbb{Q}}[\zeta^2]\Sigma, \qquad \forall \mathbb{P} \in \mathcal{D}_{\mathsf{s}}(\mu,\Sigma).$$

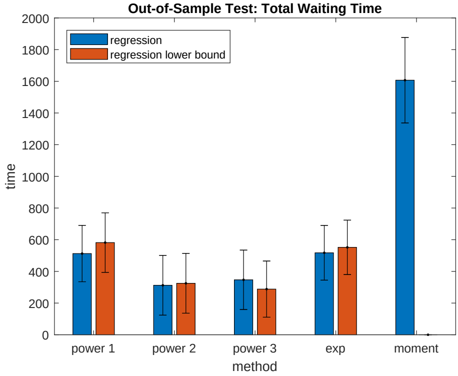
Furthermore, the upper bound $\mathbb{E}_{\mathbb{Q}}[\zeta^2]\Sigma$ is sharp in the sense that, for any symmetric matrix Δ, $\mathbb{E}_{\mathbb{P}}[(\xi-\mu)(\xi-\mu)^\top] \preceq \Delta$ for all $\mathbb{P} \in \mathcal{D}_{\mathsf{s}}(\mu,\Sigma)$ implies that $\Delta \succeq \mathbb{E}_{\mathbb{Q}}[\zeta^2]\Sigma$.

- Dominance information implies the covariance if $\mathbb{E}_{\mathbb{Q}}[\zeta^2] \leq 1$.
- Check $\mathbb{E}_{\mathbb{Q}}[\zeta^2]$ before adding in covariance info.

An Application to DR Appointment Scheduling

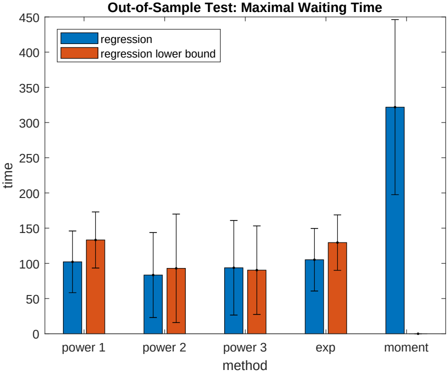
- Single server, 10 Appointments with random duration.
- Fixed sequence of arrival and scheduling arriving times.
- Mean, support, and dominance information estimated from Log-Normal samples.
  - $1 - \epsilon(r)$ obtained by fitting power and exponential curves of r.
  - Tested the fitted curve and the 95% lower envelope.
- Two objectives considered:
  - Objective 1: minimizing the total waiting time. (Efficiency)
  - Objective 2: minimizing the maximal waiting time. (Fairness)
- $||\cdot||$: $||\cdot||_\infty \Rightarrow$ LP reformulations.

Out-of-Sample Total Waiting Time



- Power$K$: regression to $\sum_{k=0}^{K} c_k r^{-k}$, $K = 1, 2, 3$.
- Exp: regression to $c_1 e^{-c_2 r}$.
- Moment: with mean and support info but without dominance info.
- Error bar: standard deviation.

Out-of-Sample Maximal Waiting Time



- Power$K$: regression to $\sum_{k=0}^{K} c_k r^{-k}$, $K = 1, 2, 3$.
- Exp: regression to $c_1 e^{-c_2 r}$.
- Moment: with mean and support info but without dominance info.
- Error bar: standard deviation.

Takeaways

- DRO approach can help address modeling and computational challenges of risk constraints.
- Structural information can...
  - make risk constraints much less conservative.
  - be incorporated without big computational burden.
- Manuscripts this talk is based on:
  - Li, B., Jiang, R., Mathieu, J.L., "Ambiguous Risk Constraints with Moment and Unimodality Information," Mathematical Programming, 2018.
  - Li, B., Jiang, R., Mathieu, J.L., "Distributionally Robust Chance-Constrained Optimal Power Flow Assuming Log-Concave Distributions," PSCC, 2018.
  - Guo, Y., Jiang, R., "Distributionally Robust Expectation Using Dominance Information," soon available, 2018.
- ruiwei@umich.edu

Thank you!