Convexity properties of stochastic programming models

Paolo Brandimarte
Dipartimento di Scienze Matematiche
Politecnico di Torino

e-mail: paolo.brandimarte@polito.it
URL: http://staff.polito.it/paolo.brandimarte

March 5, 2025

References

- P. Kall, S.W. Wallace. *Stochastic Programming*. Wiley, 1994. The book pdf can be downloaded from `http://stoprog.org/index.html?resources.html`
- J.R. Birge, F. Louveaux. *Introduction to Stochastic Programming*, (2nd ed.). Springer, 2011.
- J.L. Higle, S. Sen. *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming.* Kluwer, 1996.

Chance constrained programming

Consider the general, joint chance constraint

variabile casuale

$$\mathbb{P}\left(\{\omega \mid \mathbf{g}(\mathbf{x}, \xi(\omega)) \leq \mathbf{0}\}\right) \geq 1 - \alpha$$

where $\mathbf{g}$ is a vector-valued function.

A point $\hat{\mathbf{x}}$ is feasible if the set

$$S(\hat{\mathbf{x}}) = \{\omega \mid \mathbf{g}(\hat{\mathbf{x}}, \xi(\omega)) \leq \mathbf{0}\} \tag{1}$$

has probability measure of at least $1 - \alpha$.

e' uno spazio di probabilità

Let $\mathcal{F}$ be the field of all events and $\mathcal{G} \subset \mathcal{F}$ the collection of events with probability measure at least $1 - \alpha$. Famiglia degli eventi con prob. almeno $1-\alpha$

Then $\hat{\mathbf{x}}$ is feasible if there exists at least one event $G \in \mathcal{G}$ such that $\mathbf{g}(\hat{\mathbf{x}}, \xi(\omega)) \leq \mathbf{0}$, $\forall \omega \in G$: e' un insieme della famiglia $\mathcal{G}$

$$\hat{\mathbf{x}} \in \bigcap_{\omega \in G} \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}, \xi(\omega)) \leq \mathbf{0}\}.$$

Intersezione nello spazio delle x

CHANCE CONSTRAINED PROGRAMMING

Then, the feasible set $\mathcal{X}$ is the union of all such vectors:

$$\mathcal{X} = \bigcup_{G \in \mathcal{G}} \bigcap_{\omega \in G} \{\mathbf{x} \mid \mathbf{g}(\mathbf{x}, \boldsymbol{\xi}(\omega)) \leq \mathbf{0}\}$$

Even if each set $\{\mathbf{x} \mid \mathbf{g}(\mathbf{x}, \boldsymbol{\xi}(\omega)) \leq \mathbf{0}\}$ is convex for all $\boldsymbol{\xi}$, we should not expect convexity in general, since the union of convex sets is not guaranteed to be convex.

Convexity of the overall problem depends on the type of distribution (continuous vs. discrete) and its CDF, the kind of constraints (joint or not), and the functions in $\mathbf{g}$ (linear, convex/concave, or arbitrary).

CCP: AN EASY (NONTRIVIAL) CASE

lo assumiamo noto

Vettore di variabili casuali multivariate

Let us consider a random linear constraint $\mathbf{a}^\mathsf{T}\mathbf{x} \le b$, where $\mathbf{a} \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma)$, and require $\mathbb{P}\{\mathbf{a}^\mathsf{T}\mathbf{x} \le b\} \ge \eta$. → Come trattiamo algoritmicamente questo vincolo?

→ gioca il ruolo di $1-\alpha$

Matrice di covarianza

For a given vector $\mathbf{x}$, we have $\mathbf{a}^\mathsf{T}\mathbf{x} \sim \mathrm{N}(\nu, \sigma^2)$, where $\nu = \boldsymbol{\mu}^\mathsf{T}\mathbf{x}$ and $\sigma^2 = \mathbf{x}^\mathsf{T}\Sigma\mathbf{x}$.

Forma quadratica convessa

Using the standard normal CDF $\Phi(z)$, we rewrite the constraint as:

è il quantile di η

$$\mathbb{P}\left\{\frac{\mathbf{a}^\mathsf{T}\mathbf{x} - \nu}{\sigma} \le \frac{b-\nu}{\sigma}\right\} = \Phi\left(\frac{b-\nu}{\sigma}\right) \ge \eta \iff \frac{b-\nu}{\sigma} \ge \Phi^{-1}(\eta)$$

Cerchiamo di togliere la radice

Assuming $\eta > 0.5$, so that $\Phi^{-1}(\eta) > 0$, this is a second-order cone constraint:

Radice della matrice $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$

$$\boldsymbol{\mu}^\mathsf{T}\mathbf{x} + \Phi^{-1}(\eta)\,\|\Sigma^{1/2}\mathbf{x}\|_2 \le b,$$

→ Non è un vincolo di programmazione lineare, ma è un vincolo conico e soprattutto convesso

where $\Sigma^{1/2}$ is the square root of the covariance matrix.

Hence, an LP with *disjoint* chance constraints of this kind is a convex SOCP.

CONVEXITY OF THE RECOURSE FUNCTION

1. Questo valore atteso e' definito?

Let us consider the recourse function $\mathcal{Q}(\mathbf{x}) \equiv \mathbb{E}\left[Q(\mathbf{x}, \tilde{\xi})\right]$ in the case of deterministic recourse cost:

$$Q(\mathbf{x}, \xi) \equiv \min_{\mathbf{y}} \left\{ \mathbf{q}^{\mathsf{T}}\mathbf{y} \mid \mathbf{W}\mathbf{y} = \mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}, \mathbf{y} \geq \mathbf{0} \right\}$$

→ Valore atteso della minimizzazione

The effective domain of $\mathcal{Q}(\mathbf{x})$ consists of vectors $\mathbf{x}$ that are feasible for the first-stage constraints and such that $\mathcal{Q}(\mathbf{x}) < +\infty$, i.e., the second stage problem is (almost surely) feasible.

↳ Vogliamo vedere soluzioni ammissibili !.

Usa la dualità forte

Note: we assume that, barring second-stage infeasibility, the above expectation is always defined, i.e., random variables are not heavy-tailed.

By LP duality, we have $Q(\mathbf{x}, \xi) = \max_{\boldsymbol{\pi}} \left\{ [\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}]^{\mathsf{T}}\boldsymbol{\pi} \mid \mathbf{W}^{\mathsf{T}}\boldsymbol{\pi} \leq \mathbf{q} \right\}$. Dualizzazione del problema

Let us denote the feasible set of the dual by $\Pi = \{\boldsymbol{\pi} \mid \mathbf{W}^{\mathsf{T}}\boldsymbol{\pi} \leq \mathbf{q}\}$. Note that, since $\mathbf{q}$ is deterministic, that this set is nonrandom.

↳ Non e' stocastica!

Condizione del simplesso primario

Siamo sicuri che sia ammissibile?

Questa regione non dipende dalle ξ, quindi se hai già una la hai sempre

Convexity of the recourse function

combinazione convessa

Let us consider $\mathbf{x}^1, \mathbf{x}^2 \in \mathcal{D}$, $\lambda \in [0,1]$, and $\mathbf{x}_\lambda = \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2$.

Vogliamo dimostrare che $\mathcal{Q}(x_\lambda)$ è convessa

$\max_{x \in S} \left(f(x) + g(x) \right) \le \max_{x \in S} f(x) + \max_{x \in S} g(x)$

$\Xi \to \xi$ maiuscola

f obiettivo del problema duale

$$
\begin{aligned}
\mathcal{Q}(\mathbf{x}_\lambda) &\overset{\text{def.}}{=} \int_\Xi Q(\mathbf{x}_\lambda, \xi)\, dP \\
&= \int_\Xi \max_{\boldsymbol{\pi} \in \Pi} \left\{ [\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}_\lambda]^{\mathsf{T}} \boldsymbol{\pi} \right\} dP \\
&= \int_\Xi \max_{\boldsymbol{\pi} \in \Pi} \left\{ \lambda \left[\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}^1 \right]^{\mathsf{T}} \boldsymbol{\pi} + (1-\lambda) \left[\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}^2 \right]^{\mathsf{T}} \boldsymbol{\pi} \right\} dP \\
&\le \lambda \int_\Xi \max_{\boldsymbol{\pi} \in \Pi} \left\{ \left[\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}^1 \right]^{\mathsf{T}} \boldsymbol{\pi} \right\} dP \\
&\qquad + (1-\lambda) \int_\Xi \max_{\boldsymbol{\pi} \in \Pi} \left\{ \left[\mathbf{h}(\xi) - \mathbf{T}(\xi)\mathbf{x}^2 \right]^{\mathsf{T}} \boldsymbol{\pi} \right\} dP \\
&= \lambda \mathcal{Q}(\mathbf{x}^1) + (1-\lambda)\mathcal{Q}(\mathbf{x}^2)
\end{aligned}
$$

max della somma
$\le$
somma dei max

Convexity of the recourse function

The previous proof is rather simple, because the dual feasible region is nonrandom. Actually, convexity of the recourse function can be shown in a more general setting.

To cut a long story short (and cutting a few technical corners along the way...):

- The recourse function is typically differentiable ~~continuous~~ for continuous probability distributions.
- The recourse function is polyhedral for discrete probability distributions.

In both cases, we may rely on Kelley's cutting planes to solve the problem.

The bonus, given problem structure, is that we may find cutting ↳ iperpiani di supporto planes by a decomposition strategy.

Nel caso non differenziabile ci sono ∞ piani di supporto, ma a noi ne basta uno