Subdifferential characterization of probability functions under Gaussian distributions

A. Hantoute

CMM, Universidad de Chile

03 2017

CMM, Santiago, 2017

Joint work with R. Henrion and P. Perez-Aros

Preliminaries-probability functions

Problem

We are interested in variational properties of Gaussian probability functions induced by non-necessarily smooth initial data, $\varphi : X \to \mathbb{R}$, defined as

$$\varphi(x) := \mathbb{P}\left(g\left(x, \xi\right) \leq 0\right)$$

- where X is a reflexive and separable Banach space
- $g : X \times \mathbb{R}^m \to \mathbb{R}$ is a function depending on the realizations of an m-dimensional random vector ξ
- g is locally Lipschitzian en (x, z) and convex in z
- mainly, $\xi \sim \mathcal{N}\left(0, R\right)$
- At the reference point $\bar{x}$, $g\left(\bar{x}, 0\right) < 0$; this implies that φ is continuous at any point $x \in X$ with $g(x, 0) < 0$.

Preliminaries-probability functions

Example: Define $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by

$$g(x, z_1, z_2) := \alpha(x) e^{h(z_1)} + z_2 - 1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0, \end{cases}$$

$$h(t) := -1 - 4 \log(1 - \Phi(t)); \quad \Phi(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\tau^2/2} d\tau,$$

i.e., Φ is the distribution function of the one-dimensional standard normal distribution.

Fact

If $\xi = (\xi_1, \xi_2) \sim \mathcal{N}(0, 1)$, then g is C^1, but φ fails to be locally Lipschitzian in 0

Preliminaries - notions from variational analysis

- The Fréchet subdifferential of φ at $\bar{x}$ has the form

$$\widehat{\partial}\varphi\left(\bar{x}\right) := \left\{ x^* \in X^* \mid \liminf_{x \to \bar{x}} \frac{\varphi(x) - \varphi\left(\bar{x}\right) - \langle x^*, x - \bar{x}\rangle}{\|x - \bar{x}\|} \geq 0 \right\}$$

- The Mordukhovich subdifferential is

$$\partial\varphi\left(\bar{x}\right) := \left\{ x^* := w\text{-}\lim x_n^* \in \widehat{\partial}\varphi(x_n),\ x_n \to_{\varphi} x \right\}$$

- The singular subdifferential is

$$\partial^{\infty}\varphi\left(\bar{x}\right) := \left\{ x^* := w\text{-}\lim \lambda_n x_n^*,\ x_n^* \in \widehat{\partial}\varphi(x_n),\ x_n \to x,\ \lambda_n \searrow 0 \right\}$$

- The Clarke subdifferential is

$$\partial_C\varphi\left(\bar{x}\right) = \overline{\mathrm{co}}\ \left\{ \partial\varphi\left(\bar{x}\right) + \partial^{\infty}\varphi\left(\bar{x}\right) \right\}$$

- If φ happens to be convex, these coincide with the convex subdifferential ($\varepsilon = 0$):

$$\partial_{\varepsilon}\varphi\left(\bar{x}\right) := \left\{ x^* \in X^* \mid \varphi(x) \geq \varphi(\bar{x}) + \langle x^*, x - \bar{x}\rangle - \varepsilon,\ \forall x \in X \right\}.$$

Preliminaries

Definition

We formulate a stochastic programming problem as (Charnes, Cooper and Symonds (1958))

$$\min_{P(g_1(x,\xi)\geq 0,\cdots,g_k(x,\xi)\geq 0)\geq p} h(x)$$

or, equivalently,

$$\min_{P(\max_{i=1,k} g_i(x,\xi)\geq 0)\geq p} h(x)$$

- Models optimization problems involving uncertainties (demand, production, and so on)
- The probability, or safety, p may reflect the reliability of the system
- For p closed to 1, one ensures that the state of the system remains within a subset of all possible states under the absence of major failures.

Preliminaries

Fact

The use of the alternative formulations

$$\min_{P(g_1(x,\xi)\geq 0)\geq p_1,\cdots,P(g_k(x,\xi)\geq 0)\geq p_k} h(x)$$

may or may not be justified from the point of view of model construction; for instance, if the randome variables $g_1(x,\xi),\cdots,g_k(x,\xi)$ *are independent of each other, then*

$$P(g_1(x,\xi)\geq 0,\cdots,g_k(x,\xi)\geq 0) = P(g_1(x,\xi)\geq 0).\dots.P(g_k(x,\xi)\geq 0)$$

Preliminaries-Spheric-radial decomposition of Gaussian random vectors

Fact

For any Borel set $M \subset \mathbb{R}^m$

$$\mathbb{P}\left(\xi \in M\right) = \int\limits_{v \in \mathbb{S}^{m-1}} \mu_\eta\left(\{r \geq 0 \mid rLv \in M\}\right) d\mu_\zeta(v),$$

- μ_η is the one-dimensional Chi-distribution with m degrees of freedom
- μ_ζ is the uniform distribution on $\mathbb{S}^{m-1}$
- L is a factor in a decomposition $R = LL^T$

Fact

Hence, defining the radial probability function $e : X \times \mathbb{S}^{m-1} \to \mathbb{R}$ as $e(x, v) := \mu_\eta\left(\{r \geq 0 \mid g(x, rLv) \leq 0\}\right)$,

$$\varphi(x) = \int\limits_{\mathbb{S}^{m-1}} e(x, v) d\mu_\zeta(v)$$

Preliminaries - Explicit expression of probability functions

For $x \in X$ satisfying $g(x,0) < 0$, we set

$$F(x) = \{v \in \mathbb{S}^{m-1} \mid \exists r \geq 0 : g(x, rLv) = 0\}$$
$$I(x) = \{v \in \mathbb{S}^{m-1} \mid \forall r \geq 0 : g(x, rLv) < 0\}$$

and

$$\rho(x,v) := \begin{cases} r \text{ such that } g(x, rLv) = 0 & \text{if } v \in F(x) \\ +\infty & \text{if } v \in I(x), \end{cases}$$

Fact

We get

$$\varphi(x) = \int_{\mathbb{S}^{m-1}} F_\eta(\rho(x,v))\, d\mu_\zeta(v),$$

where F_η is the distribution function of the Chi-distribution with m degrees of freedom

$$F'_\eta(t) = \chi(t) := Kt^{m-1}e^{-t^2/2} \quad \forall t \geq 0.$$

Preliminaries - Continuity of probability functions

Lemma

Define $U := \{x \in X \mid g(x,0) < 0\}$.

1. *The radius function* ρ *is continuous at* (x, v) *for any* $x \in U$ *and any* $v \in F(x)$.
2. *For* $x \in U$ *and* $v \in I(x)$ *it holds that* $\lim_{k\to\infty} \rho(x_k, v_k) = \infty$ *for any sequence* $(x_k, v_k) \to (x, v)$ *such that* $v_k \in F(x_k)$.

Lemma

The radial probability function $e(x, v) = F_\eta(\rho(x, v))$ *is continuous at any* $(x, v) \in X \times \mathbb{S}^{m-1}$ *with* $g(x, 0) < 0$.

Theorem

The probability function is continuous at any point $x \in X$ *with* $g(x, 0) < 0$.

Normal Integrands

We consider a normal integrand

$$f : T \times X \to \overline{\mathbb{R}}$$

- (T, Σ, μ) with σ-finite, complete and positive measure
- f is a $\Sigma \otimes \mathcal{B}(X)$-measurable and $f(t, \cdot)$ is lsc
- If in addition $f(t, \cdot) \in \Gamma(X)$ then f is convex normal integrand

Definition

We introduce the mapping $I_f : X \to \overline{\mathbb{R}}$

$$I_f(x) := \int_T f(t, x) d\mu(t)$$

Normal Integrands

- g is Gelfand integrable if for each $A \in \Sigma$

$$\langle x_A^*, \cdot \rangle = \int_A \langle g(t), \cdot \rangle \, d\mu(t) \in X^*$$

Definition

The Gelfand integral of multifunction $M : T \rightrightarrows X^*$ over $A \in \sum$ is the set

$$\int_A M(t) d\mu(t) := \left\{ \int_A x^*(t) d\mu(t) \mid x^*(\cdot) \text{ is a G-integrable selector of } M \right\}$$

Convex normal Integrands

Theorem

If $f(t,x) \geq \langle \gamma(t), x\rangle + \beta$, $\gamma \in L^\infty(T,X)$, $\beta \in L^\infty(T)$, *then*

$$\partial I_f(x) = \bigcap_{\substack{\varepsilon>0 \\ F\in\mathcal{F}(x)}} \mathrm{cl}^{w^*} \int_T \left(\partial_\varepsilon f_t(x) + \mathrm{N}_{F\cap \mathrm{dom}\, I_f}(x)\right) d\mu(t)$$

($\mathcal{F}(x) = \{$*finite-dimensional subspaces* $X \supset F \ni x\}$*)*

- Ex: $I_f(x) = \int_0^1 \frac{x^2}{t} dt$ so that $\partial I_f(0) = \mathbb{R}$ and

$$\bigcap_{\varepsilon>0} \mathrm{cl}^{w^*} \left(\int_0^1 \partial_\varepsilon f_t(0) d\mu(t)\right) = \bigcap_{\varepsilon>0} \bigcup_{n\geq 1} [-4\sqrt{\varepsilon}, 4\sqrt{\varepsilon}] = \{0\}$$

- To compare with

$$\partial I_f(x) = \bigcap_{\varepsilon>0} \mathrm{cl}^{w^*} \left(\bigcup_{\varepsilon(\cdot)\in\mathcal{L}_\varepsilon} \int_0^1 \partial_{\varepsilon(t)} f_t(0) d\mu(t)\right)$$

Convex normal Integrands

- If $\widetilde{I}_f(x(\cdot)) = \int_T f(t, x(t))d\mu(t)$ is continuous at some point in $L^\infty(T, X)$ (Rockafellar 70')

$$\partial I_f(x) = \bigcup_{\varepsilon(\cdot)\in\mathcal{L}_\varepsilon} \int_T \partial f(t, x),$$

- Assuming $X = \mathbb{R}^n$ and $\operatorname{dom} I_f \cap \operatorname{ri}(\operatorname{dom} f(t, \cdot)) \neq \emptyset$ (Ioffe 2006)

$$\partial I_f(x) = \int_T (\partial f(t, x) + \mathrm{N}_{\operatorname{dom} I_f}(x))d\mu$$

- $\widetilde{I}_f : L^\infty(T, X) \to \overline{\mathbb{R}}$ norm-continuous at some constant point, then (Ioffe-Levin 70')

$$\partial I_f(x) = \int_T \partial f(t, x)d\mu + \mathrm{N}_{\operatorname{dom} I_f}(x)$$

Convex normal Integrands

- (Ioffe (2006) and Thibault-Lopez (2008))

$$\partial I_f(x) = \lim_i \int_T \partial f(t, x_i(t)), \quad \int_T f(t, x_i(t)) \to \int_T f(t, x)$$

- For $g, h \in \Gamma_0(X)$

$$\begin{aligned} \partial(g+h)(x) &= \bigcap_{\substack{\varepsilon>0 \\ F\in\mathcal{F}(x)}} \mathrm{cl}^{w^*}\left(\partial_\varepsilon g(x) + \partial_\varepsilon h(x) + \mathrm{N}_{F\cap \mathrm{dom}\, g \cap \mathrm{dom}\, h}(x)\right) \\ &= \bigcap_{\varepsilon>0} \mathrm{cl}^{w^*}\left(\partial_\varepsilon g(x) + \partial_\varepsilon h(x)\right) \end{aligned}$$

Lsc normal integrands

We assume that $I_f(x_0) < +\infty$ for some $x_0 \in X$ and that for some $\delta > 0$ and $K \in L^1(T, \mathbb{R})$ we have

$$\widehat{\partial}_x f(t,x) \subseteq K(t) B_1^*(0) + C, \ \forall x \in B_\delta(x_0), \ ae \ t \in T, \tag{1}$$

where $C \subseteq X^*$ is a closed convex cone with polar cone having a nonempty interior.

Theorem

We have that

$$\partial I_f(x_0) \subseteq \mathrm{cl}^* \left\{ \int_T \partial f(t, x_0) d\mu(t) + C \right\}, \quad \partial^\infty I_f(x_0) \subseteq C,$$

$$\partial_C I_f(x_0) \subseteq \overline{\mathrm{co}} \left\{ \int_T \partial f(t, x_0) d\mu(t) + C \right\}.$$

A kind of Mean Value Theorem

Let $x \in X$ with $g(x,0) < 0$ and $v \in F(x)$ be arbitrary.

Fact

$\forall y^* \in \widehat{\partial}_x e(x,v)$ and $\forall w \in X$, $\exists$ $(x^*, z^*) \in \partial_C g(x, \rho(x,v)Lv)$ such that

$$\langle y^*, w \rangle \leq \frac{-F'_\eta(\rho(x,v))}{\langle z^*, Lv \rangle} \langle x^*, w \rangle .$$

Consequently, (i) There exist neighborhoods $\tilde{U}$ of x, $\tilde{V}$ of v, and $\alpha > 0$ such that

$$\widehat{\partial}_x e(x', v') \subseteq \mathbb{B}^*_\alpha(0) \quad \forall (x', v') \in \tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1} \right).$$

Fact

(ii) For all $x \in X$ with $g(x,0) < 0$ and for all $v \in I(x)$ one has that

$$\widehat{\partial}_x e(x,v) \subseteq \{0\}.$$

Cone of nice directions

Definition

For $x \in X$ and $l > 0$, the *l-cone of nice directions at* $x \in X$ is

$$C_l(x) := \{h \in X \mid g^\circ(\cdot, z)(y; h) \leq l \|z\|^{-m} e^{\frac{\|z\|^2}{2\|L\|^2}} \|h\| \ \forall y \in \mathbb{B}_{1/l}(x), \|z\| \geq l\},$$

where $g^\circ(\cdot, z)$ is the Clarke derivative

Fact

Fix $x_0 \in X$ *such that* $g(x_0, 0) < 0$. *Then, for every* $l > 0$, *there exists some neighborhood* U *of* x_0 *and some* $R > 0$ *such that*

$$\partial_x^F e(x, v) \subseteq \mathbb{B}_R^*(0) - C_l^*(x_0) \quad \forall x \in U,\ v \in \mathbb{S}^{m-1}.$$

Subdifferential of the Gaussian probability function

Theorem

Let $x_0 \in X$ be such that $g(x_0, 0) < 0$. Assume that the cone $C_l(x_0)$ has a non-empty interior for some $l > 0$. Then,

- $\partial\varphi(x_0) \subseteq \mathrm{cl}^* \left\{ \int\limits_{v \in \mathbb{S}^{m-1}} \partial_x e(x_0, v) d\mu_\zeta(v) - C_l^*(x_0) \right\}$
- $\partial^\infty \varphi(x_0) \subseteq -C_l^*(x_0)$
- $\partial_C \varphi(x_0) \subseteq \overline{\mathrm{co}} \left\{ \int\limits_{v \in \mathbb{S}^{m-1}} \partial_x e(x_0, v) d\mu_\zeta(v) - C_l^*(x_0) \right\}$

Subdifferential of the Gaussian probability function

Corollary

Moreover, suppose that $\partial_x e(x_0, v)$ is integrably bounded; i.e., there exists some integrable function $K : \mathbb{S}^{m-1} \to \mathbb{R}_+$ such that

$$\partial_x e(x_0, v) \subseteq \mathbb{B}^*_{K(v)}(0) \quad \mu_\zeta - a.e.\ v \in \mathbb{S}^{m-1}.$$

Then

$$\partial \varphi(x_0) \subseteq \partial_C \varphi(x_0) \subseteq \mathrm{cl} \left\{ \int_{v \in \mathbb{S}^{m-1}} \partial_x e(x_0, v) d\mu_\zeta(v) \right\} - C^*_\infty(x_0).$$

Lipschitzianity of the Gaussian probability function

Corollary

Fix $x \in X$ such that $g(x,0) < 0$. Under one of the alternative conditions
(i) $\{z \in \mathbb{R}^m \mid g(x_0, z) \leq 0\}$ is a bounded set
(ii) $\exists\ l > 0$ such that $C_l(x_0) = X$,
the probability function φ is locally Lipschitz near x and

$$\partial_C \varphi(x) \subseteq \int_{v \in \mathbb{S}^{m-1}} \partial_x^C e(x,v) d\mu_\zeta(v).$$

Corollary

In addition if $\#\partial_x^C e(x,v) = 1 \quad \mu_\zeta$-a.e. $v \in \mathbb{S}^{m-1}$, then φ is strictly differentiable at x and

$$\nabla \varphi(x) = \int_{v \in \mathbb{S}^{m-1}} \nabla_x e(x,v) d\mu_\zeta(v);$$

thus, φ is C^1 if X is finite-dimensional.

Application to a finite system of smooth inequalities

Problem

We introduce

$$\varphi(x) := \mathbb{P}\left(g_i\left(x, \xi\right) \leq 0,\ i = 1, \ldots, p\right),\ x \in X$$

for g_i Lipschitz.

Application to a finite system of smooth inequalities

Fact

If $g := \max_{i=1,\dots,p} g_i$*, then we go back to our setting where*

$$\rho(x,v) = \min_{i=1,\dots,p} \rho_i(x,v) \quad \forall x : g(x,0) < 0,\ \forall v \in F(x).$$

The functions ρ_i *are* C^1 *and for all* x *with* $g(x,0) < 0$ *and all* $v \in F(x)$*,*

$$\nabla_x \rho_i(x,v) = -\frac{1}{\langle \nabla_z g_i(x, \rho(x,v) Lv), Lv \rangle} \nabla_x g_i(x, \rho(x,v) Lv),\ \ i = 1, \dots, p.$$

Application to a finite system of smooth inequalities

Theorem

Fix $x_0 \in X$ with $g(x_0, 0) < 0$, and assume that for some $l > 0$ it holds

$$\|\nabla_x g_i(x,z)\| \leq l \, \|z\|^{-m} \, e^{\frac{\|z\|^2}{2\|L\|^2}} \quad \forall x \in \mathbb{B}_{1/l}(x_0), \; \|z\| \geq l.$$

Then φ is locally Lipschitz near x_0 and there exists some $R > 0$ such that

$$\partial_C \varphi(x_0) \subseteq -\int_{v \in F(x_0)} \operatorname{co} \left\{ \bigcup_{i \in T(v)} \frac{\chi(\rho(x_0, v)) \nabla_x g_i(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_i(x_0, \rho(x_0, v) Lv), Lv \rangle} \right\} d\mu_\zeta(v)$$

$$+\mu_\zeta(I(x_0)) \mathbb{B}^*_R(0).$$

References

- W. Van Ackooij and R. Henrion, *Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions*. SIAM J. Optim. (2014).
- W. Van Ackooij and R. Henrion, *(Sub-) Gradient formulae for probability functions of random inequality systems under Gaussian distribution*. SIAM/ASA J. Uncertainty Quantification (2017).
- C. Castaing and M. Valadier, Convex Analysis and Measurable Multifunctions, 1977.
- AD. Ioffe and V.L. Levin, Subdifferentials of convex functions, Trans. Moscow Math. Soc. 26 (1972).
- AD Ioffe, Three Theorem on convex integrals, J. Conv. Analysis, (2006)
- C. Gotzes, H. Heitsch, R. Henrion and R. Schultz, *Feasibility of nominations in stationary gas networks with random load*, Mathematical Methods of Operations Research (2016).

References

- B. S. Mordukhovich and N. Sagara, Differentials of nonconvex integral functionals in Banach spaces with applications to stochastic dynamic programming, JCA (2016)
- O. Lopez, L. Thibault, J. Convex and Nonliner Analysis (2011)
- A. Prékopa, Stochastic Programming, Kluwer, Dordrecht, 1995.
- RT. Rockafellar, Integrals which are convex functionals. I, II, Pacific J. Math. (1968, 1971).
- RT. Rockafellar, Existence and duality theorems for convex problems of Bolza, Trans. Amer. Math. Soc. (1971).

Thank you!