Weierstrass Institute for
Applied Analysis and Stochastics

# Subdifferential Characterization of Gaussian probability functions

R. Henrion

Weierstrass Institute Berlin

Joint work with A. Hantoute, P. Perez Aros (CMM, Santiago)

## Probability functions

We consider probability functions of the type

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

where

- $x \in X$ is a decision variable in a separable and reflexive Banach space $X$
- $\xi$ is an $m$-dimensional random vector defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$
- $g : X \times \mathbb{R}^m \to \mathbb{R}$ is a mapping defining the random inequality constraint $g(x,\xi) \leq 0$

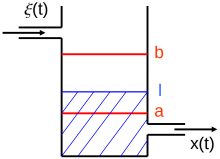
Our basic assumptions:

- $g$ locally Lipschitzian
- $g(x,\cdot)$ convex for all $x \in X$
- $\xi$ is a Gaussian random vector

Probability functions occur in many optimization problems from engineering, e.g.

$$\max\{\varphi(x) \mid x \in X\} \quad \text{reliability maximization}$$
$$\min\{f(x) \mid \varphi(x) \geq p\} \quad \text{probabilistic constraints}$$

## Reservoir control problem




Consider a reservoir with random inflow $\xi$ and controlled release $x$:

Assume a finitely parameterized inflow process

$\xi(t) = \langle \xi, a(t) \rangle, \quad \xi \sim \mathcal{N}(\mu, \Sigma)$ (e.g., K-L expansion)

Water level at time $t$:

$$l(\xi, x, t) = l_0 + \int_0^t \langle \xi, a(\tau) \rangle d\tau - \int_0^t x(\tau) d\tau$$

Probability of satisfying a critical lower level profile $l_*$ given a release profile $x$:

$$\varphi(x) := \mathbb{P}(l(\xi, x, t) \geq l_*(t) \quad \forall t \in [0, T]) = \mathbb{P}\left( \underbrace{\max_{t \in [0,T]} \{l_*(t) - l(\xi, x, t)\}}_{g(x,\xi)} \leq 0 \right)$$

$g$ locally Lipschitz and convex in $\xi \implies$ basic assumptions satisfied.

## Slater point assumption

Let $\bar{x} \in X$ be a point of interest for our probability function $\varphi(x) := \mathbb{P}(g(x,\xi) \le 0)$.

In addition to our basic assumptions

$g$ locally Lipschitz, $g(x,\cdot)$ convex, $\xi \sim \mathcal{N}(\mu,\Sigma)$

suppose that: $g(\bar{x},\mu) < 0$ (mean is a Slater point).

Slater point assumption

- is satisfied whenever $\varphi(\bar{x}) \ge 0.5$ $\Longrightarrow$ no restriction of generality
- implies continuity of $\varphi$ at $\bar{x}$.

Question: Does the Slater point assumption for the mean along with $g \in \mathcal{C}^1$ imply that $\varphi \in \mathcal{C}^1$?

Answer: **No** in general, **Yes** for $g$ linear in $\xi$.

## Possibly Non-Lipschitzian $\varphi(x) = \mathbb{P}(g(x,\xi) \le 0)$ for $g \in \mathcal{C}^1$

Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ and

$$g(x,z) := \langle a(x), z\rangle - b(x), \quad a \in \mathcal{C}^1(X, \mathbb{R}^m),\ b \in \mathcal{C}^1(X, \mathbb{R}),\ X \text{ - Banach space}$$

Slater point assumption at point of interest: $\langle a(\bar{x}), \mu\rangle < b(\bar{x})$. Then, with $\Phi$ = CDF of $\mathcal{N}(0,1)$:

$$\varphi(\bar{x}) = \Phi\left(\frac{b(\bar{x}) - \langle a(\bar{x}), \mu\rangle}{\sqrt{\langle a(\bar{x}), \Sigma a(\bar{x})\rangle}}\right) \in \mathcal{C}^1$$

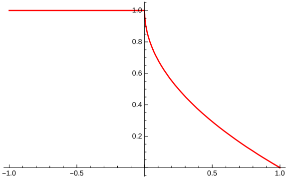
---

Let $g(x, z_1, z_2) := x^2 \cdot 1_{[0,\infty)}(x) \cdot \exp(-1 - 4\log(1 - \Phi(z_1))) + z_2 - 1 \in \mathcal{C}^1$.

Then, $g$ is convex in $(z_1, z_2)$ for every $x \in \mathbb{R}$.

Let $\xi = (\xi_1, \xi_2) \sim \mathcal{N}(0, I_2)$. Then, $g(\bar{x} := 0, \mu = 0) < 0$

(Slater point assumption) and




$$\varphi(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int e^{-s^2/2}\Phi(1)ds & x \le 0 \\ \frac{1}{\sqrt{2\pi}} \int e^{-s^2/2}\Phi(1 - x^2\exp(-1 - 4\log(1 - \Phi(s))))ds & x > 0 \end{cases}$$

$\varphi$ is continuous (by Slater point assumption) but **not even locally Lipschitz**.

## A quick reminder of subdifferentials

**Definition**

Let $X$ be a Banach space and $f : X \to \mathbb{R}$ lsc. Then, the Fréchet subdifferential of $f$ at $\bar{x} \in X$ is defined as

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

If $X$ is a reflexive Banach space, then the limiting (Mordukhovich) subdifferential of $f$ at $\bar{x} \in X$ is defined as

$$\partial^M f(\bar{x}) := \left\{ x^* \in X^* \mid \exists x_n \to \bar{x}, x_n^* \rightharpoonup x^* : x_n^* \in \partial^F f(x_n) \right\}.$$

If $f$ is locally Lipschitzian, then Clarke's subdifferential is obtained from the limiting one by

$$\partial^C f(\bar{x}) = \overline{\text{co}}\, \partial^M f(\bar{x})$$

Example: $\partial^F(-|\cdot|)(0) = \emptyset,\ \partial^M(-|\cdot|)(0) = \{-1, 1\},\ \partial^C(-|\cdot|)(0) = [-1, 1].$

## Spheric-radial decomposition of a Gaussian random vector

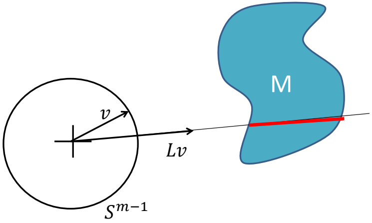
Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ with $\Sigma = LL^T$. Then,

$$\mathbb{P}(\xi \in M) = \int_{v \in \mathbb{S}^{m-1}} \mu_\eta\left(\{r \geq 0 : \mu + rLv \cap M \neq \emptyset\}\right) d\mu_\zeta(v),$$

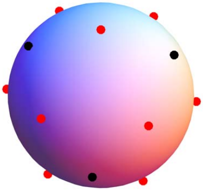
where $\mu_\eta, \mu_\zeta$ are the laws of $\eta \sim \chi(m)$ and of the uniform distribution on $\mathbb{S}^{m-1}$.

For a parameter-dependent set:

$$\varphi(x) = \mathbb{P}(g(x,\xi) \leq 0) = \int_{v \in \mathbb{S}^{m-1}} \underbrace{\mu_\eta\left(\{r \geq 0 : g(x, \mu + rLv) \leq 0\}\right)}_{e(x,v):\ \text{radial probability function}} d\mu_\zeta(v),$$




QMCsampling of the sphere



## The cone of nice directions

### Definition

According to our basic assumptions, let $g : X \times \mathbb{R}^m \to \mathbb{R}$ be locally Lipschitz.
For $l > 0$, we define the l- cone of nice directions at $\bar{x} \in \mathbb{R}^n$, as

$$C_l \quad := \quad \Big\{ h \in X \mid d^C g(\cdot, z)(x; h) \le l \|z\|^{-m} \exp(\|z\|^2/(2\|L\|^2)) \|h\| \\ \forall x \in \mathbb{B}_{1/l}(\bar{x}) \ \forall z : \|z\| \ge l \Big\}$$

Here (Clarke's directional derivative of partial function),

$$d^C g(\cdot, z)(x; h) := \limsup_{y \to x,\, t \downarrow 0} \frac{g(y + th, z) - g(y, z)}{t}$$

If $g \in \mathcal{C}^1$, then $d^C g(\cdot, z)(x; h) = \langle \nabla_x g(x, z), h \rangle = g'(\cdot, z)(x; h)$.

### Proposition

*Let $\bar{x} \in X$ such that $g(\bar{x}, \mu) < 0$. Then, for every $l > 0$ there exists a neighbourhood $U$ of $\bar{x}$ such that*

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

**Theorem (Correa, Hantoute, Perez-Aros (2016))**

*Let $(\Omega, \mathcal{A}, \nu)$ a $\sigma$- finite measure space and $f : \Omega \times X \to [0, \infty]$ a normal integrand. Define the integral functional*

$$I_f(x) := \int_{\omega \in \Omega} f(\omega, x) d\nu.$$

*Assume that for some $\delta > 0$, $K \in L^1(\Omega, \mathbb{R})$ and some closed cone $C \subseteq X$ having nonempty interior:*

$$\partial_x^F f(\omega, x) \subseteq K(\omega)\mathbb{B}_1^*(0) + C^* \quad \forall x \in \mathbb{B}_\delta(x_0)\ \forall \omega \in \Omega.$$

*Then,*

$$\partial^M I_f(x_0) \subseteq \mathrm{cl}^* \left\{ \int_{\omega \in \Omega} \partial^M f(\omega, x^0) d\nu(\omega) + C^* \right\}$$

### Theorem (Hantoute, H., Pérez-Aros 2017)

*Assume that $g : X \times \mathbb{R}^m \to \mathbb{R}$ is locally Lipschitz and convex in the second argument. Moreover, let $\xi \sim \mathcal{N}(\mu, \Sigma)$ and fix a point $\bar{x}$ satisfying $g(\bar{x}, \mu) < 0$. Finally, suppose that for some $l > 0$ the l-cone $C_l$ of nice directions at $\bar{x}$ has nonempty interior. Then,*

$$\partial^M \varphi(\bar{x}) \subseteq \mathrm{cl}^* \left\{ \int_{v \in \mathbb{S}^{m-1}} \partial_x^M e(\bar{x}, v) d\mu_\zeta(v) - C_l^* \right\}$$

*Here, $\partial^M$ refer to the Mordukhovich subdifferential, $\mu_\zeta$ is the uniform distribution on $\mathbb{S}^{m-1}$ and*

$$e(x,v) := \mu_\eta\{r \geq 0 \mid g(x, \mu + rLv) \leq 0\}, \quad (x,v) \in X \times \mathbb{S}^{m-1}; \quad (LL^T = \Sigma),$$

*where $\mu_\eta$ is the $\chi$-distribution with $m$ degrees of freedom.*

### Example

In the non-differentiable example before, we have (for $l > 0$ large enough) that

$$\partial^M \varphi(\bar{x}) = \{0\},\ C_l = (-\infty, 0],\ \partial_x^M e(\bar{x}, v) = \{0\} \text{ for } \mu_\zeta - a.e.\ v,$$

whence the inclusion in the Theorem reads here as: $\{0\} \subseteq (-\infty, 0]$.

**Theorem (Hantoute, H., Pérez-Aros 2017)**

*Assume that $g : X \times \mathbb{R}^m \to \mathbb{R}$ is locally Lipschitz and convex in the second argument. Moreover, let $\xi \sim \mathcal{N}(\mu, \Sigma)$ and fix a point $\bar{x}$ satisfying $g(\bar{x}, \mu) < 0$. Finally, suppose that $C_l = X$ for some $l > 0$ or that the set $\{z \mid g(\bar{x}, z) \leq 0\}$ is bounded. Then, $\varphi$ is locally Lipschitzian around $\bar{x}$ and*

$$\partial^C \varphi(\bar{x}) \subseteq \int_{v \in \mathbb{S}^{m-1}} \partial_x^C e(\bar{x}, v) d\mu_\zeta(v); \quad (\partial^C = \textit{Clarke subdifferential}).$$

For locally Lipschitzian functions $f$ one always has that $\emptyset \neq \partial^C f(\bar{x})$ and

$$\#\partial^C f(\bar{x}) = 1 \iff f \text{ strictly differentiable at } \bar{x}$$

**Corollary**

*In addition to the assumptions above, assume that $\#\partial_x^C e(\bar{x}, v) = 1$ for $\mu_\zeta$-a.e. $v$. Then, $\varphi$ is strictly differentiable at $\bar{x}$ and*

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

## Partial (Clarke-) subdifferential of $e(x,v)$

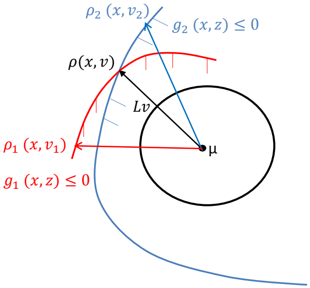
**Theorem (v. Ackooij / H. 2015)**

*For* $g(x,z) := \max_{i=1,\ldots,p} g_i(x,z)$ *and* $\xi \sim \mathcal{N}(\mu,\Sigma)$ *suppose that*

- $g_i \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m)$ *and convex in the second argument*
- $C = \mathbb{R}^n$ *(all directions nice);* $g_i(\bar{x},\mu) < 0$ *for* $i = 1,\ldots,n$ *(Slater point)*

*Then,* $\partial_x^C e(\bar{x},v) = Co\left\{-\frac{\chi(\rho(\bar{x},v))}{\langle \nabla_z g_i(\bar{x},\rho(\bar{x},v)Lv), Lv\rangle} \nabla_x g_i(\bar{x},\rho(\bar{x},v)Lv) : i \in I(v)\right\}$

*Here,* $I(v) := \{i \mid \rho(\bar{x},v) = \rho_i(\bar{x},v)\}$ *and* $\chi$ *is the density of the Chi-distribution with* $m$ *d.f.*




If $\mu_\zeta(\{v \in \mathbb{S}^{n-1} \mid \#I(v) \geq 2\}) = 0$ then

$\varphi$ is strictly differentiable at $\bar{x}$.

### Corollary

*In addition to the assumptions of the previous theorem assume the following constraint qualification:*

$$\operatorname{rank}\{\nabla_z g_i(\bar{x},z), \nabla_z g_j(\bar{x},z)\} = 2 \quad \forall i \neq j \in \mathcal{I}(z) \; \forall z : g(\bar{x},z) \leq 0,$$

*where,* $\mathcal{I}(z) := \{i \mid g_i(\bar{x},z) = 0\}$.

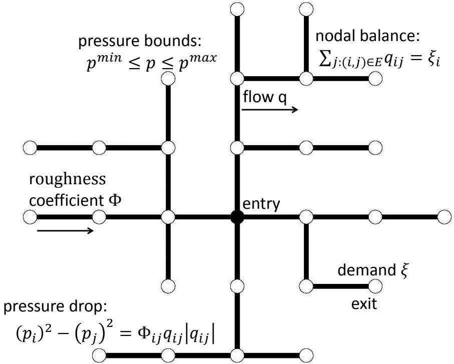
*Then,* $\varphi$ *is strictly differentiable at* $\bar{x}$. *If this condition holds locally around* $\bar{x}$, *then* $\varphi$ *is continuously differentiable. Moreover the gradient formula*

$$\nabla\varphi(\bar{x}) = -\int\limits_{v \in \mathbb{S}^{m-1}} \frac{\chi(\rho(\bar{x},v))}{\langle \nabla_z g_{i^*(v)}(\bar{x}, \rho(\bar{x},v) Lv), Lv \rangle} \nabla_x g_{i^*(v)}(\bar{x}, \rho(\bar{x},v) Lv) \, d\mu_\zeta(v)$$

*holds true. Here,* $i^*(v) := \{i | \rho(\bar{x},v) = \rho_i(\bar{x},v)\}$.

# Feasibility of random demands in a gas network

Consider a simple algebraic model of a gas network $(V, E)$:




demand vector $\xi$ feasible

$\Longleftrightarrow$

$\exists p, q$ :

$Aq = \xi,\ A^T p^2 = -\Phi |q| q,$

$p^{\min} \leq p \leq p^{\max}$

($A =$ incidence matrix)

Explicit inequality system for a tree: demand vector $\xi$ feasible $\Longleftrightarrow$ [1]

$$(p_k^{\max})^2 + g_k(\xi, \Phi) \geq (p_l^{\min})^2 + g_l(\xi, \Phi) \quad (k, l = 0, \ldots, |V|)$$

$$g_k(\xi, \Phi) = \sum_{e \in \Pi(k)} \Phi_e \left( \sum_{t \in V : t \geq h(e)} \xi_t \right)^2$$

[1] see: Gotzes, Heitsch, H. Schultz 2016

## Mixed probabilistic and robust constraint

The network owner is interested in guaranteeing the feasibility of a random demand with given probability:

$$\mathbb{P}\left((p_k^{\max})^2 + g_k(\xi, \Phi) \geq (p_l^{\min})^2 + g_l(\xi, \Phi) \quad (k, l = 0, \ldots, |V|)\right) \geq p$$

Roughness coefficient $\Phi$ uncertain too. In contrast with $\xi$ one does not have access to statistical information in general. Worst-case model with respect to a rectangular or ellipsoidal uncertainty set:

$$\begin{gathered}\mathbb{P}\big((p_k^{\max})^2 + g_k(\xi, \Phi) \geq (p_l^{\min})^2 + g_l(\xi, \Phi) \quad (k, l = 0, \ldots, |V|) \\ \forall \Phi \in [\bar{\Phi} - \delta, \bar{\Phi} + \delta] \quad \text{or:} \quad \forall \Phi : (\Phi - \bar{\Phi})^T \Sigma_\delta (\Phi - \bar{\Phi}) \leq 1\big) \geq p\end{gathered} \tag{1}$$

Here, $\bar{\Phi}$ is a nominal vector of roughness coefficients.

Infinite system of random inequalities. Mixed model of probabilistic and robust constraints.

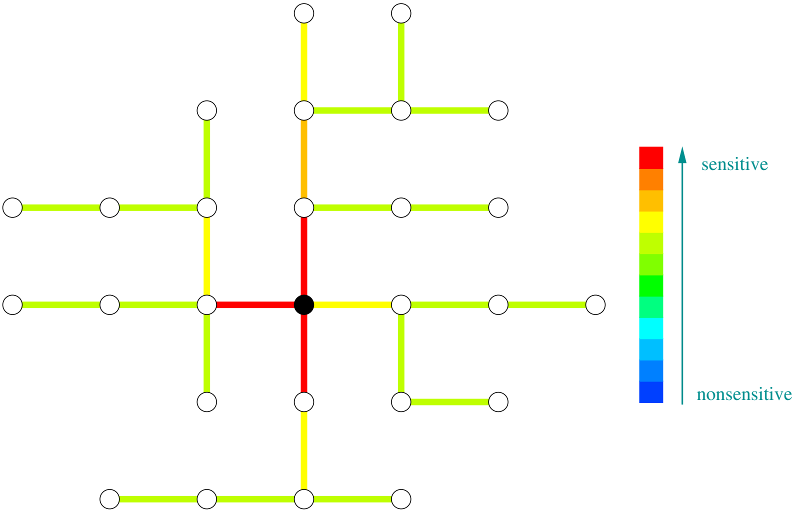
Choice of $\delta$ often not evident. In order to to gain information about local sensibility w.r.t. uncertainty in $\Phi$, we define the following optimisation problem: locale de l'incertitude en $\Phi$:

'Maximize' uncertainty set while keeping feasibility of demands with given probability:

$$\text{maximize} \quad \sum_{e \in E} \delta_e^{0.9} \quad \text{under probabilistic constraint (1)}$$

## Numerical solution for an example

Illustration of the optimal solution for a tree with 27 nodes, $p = 0.9/0.8$, $\xi$ Gaussian:



# Numerical solution for an example

Illustration of the optimal solution for a tree with 27 nodes, $p = 0.9/0.8$, $\xi$ Gaussian: