Weierstrass Institute for
Applied Analysis and Stochastics

Contraintes en probabilité - formules du gradient et applications

1ère Journée MAS-MODE

R. Henrion

Weierstrass Institute Berlin

Fonctions probabilistes

On considère des fonctions probabilistes du type $\varphi(x) := \mathbb{P}(g(x,\xi) \leq 0)$, où

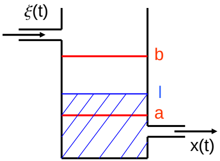
- $x \in X$ est une variable de décision dans un espace Banach séparable et réflexif
- $\xi \sim \mathcal{N}(\mu, \Sigma)$ est un vecteur Gaussien de dimension m
- $g : X \times \mathbb{R}^m \to \mathbb{R}$ est une application localement Lipschitz et convexe par rapport à l'aléa.

Les fonctions probabilistes apparaissent dans beaucoup de problème d'ingénieurerie, e.g.

$$\max\{\varphi(x) \mid x \in X\} \quad \text{maximisation de la sûreté}$$
$$\min\{f(x) \mid \varphi(x) \geq p\} \quad \text{contrainte en probabilité}$$

Problème du contrôle d'un réservoir

Considérons un réservoir avec des apports aléatoires ξ et un turbinage x:



Supposons que le processus des apports soit parametrisé par un vecteur Gaussien:

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

Niveau du réservoir à l'instant t:

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

Probabilité de satisfaire un niveau de réservoir inférieur $l_*(t)$ étant donné un turbinage 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 localement Lipschitz et convexe en $\xi \implies$ hypothèses de notre modèle satisfaites.

Fonctions probabilistes

On considère des fonctions probabilistes du type $\varphi(x) := \mathbb{P}(g(x,\xi) \leq 0)$, où

- $x \in X$ est une variable de décision dans un espace Banach séparable et réflexif
- $\xi \sim \mathcal{N}(\mu, \Sigma)$ est un vecteur Gaussien de dimension m
- $g : X \times \mathbb{R}^m \to \mathbb{R}$ est une application localement Lipschitz et convexe par rapport à l'aléa.

Même si $g \in \mathcal{C}^1$, alors φ n'est pas forcément continue:

$$\xi \sim \mathcal{N}(0,1), \quad g(x,\xi) := x \cdot \xi \leq 0, \quad \varphi(0) = 1, \quad \varphi(x) = 0.5 \ (x \neq 0)$$

Pour $\bar{x}$ fixé, on va supposer que: $g(\bar{x}, \mu) < 0$ (moyenne est un point de Slater). Cette condition

- est satisfaite si $\varphi(\bar{x}) \geq 0.5 \implies$ aucune perte de généralité
- entraine la continuité de φ en $\bar{x}$.

Question: μ point de Slater + $g \in \mathcal{C}^1 \implies \varphi \in \mathcal{C}^1$?

Réponse: **Non** en général, **Oui** pour g linéaire en ξ.

Caractère non-Lipschitz de $\varphi(x) = \mathbb{P}(g(x,\xi) \le 0)$ pour $g \in \mathcal{C}^1$

Soit $\xi \sim \mathcal{N}(\mu, \Sigma)$ et

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

Si μ est un point de Slater: $\langle a(\bar{x}), \mu\rangle < b(\bar{x})$, alors (avec $\Phi =$ CDF de $\mathcal{N}(0,1)$):

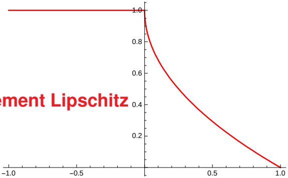
$$\varphi(\bar{x}) = \Phi\left(\frac{b(\bar{x}) - \langle a(\bar{x}), \mu\rangle}{\langle a(\bar{x}), \Sigma a(\bar{x})\rangle}\right) \in \mathcal{C}^1$$

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

Alors, g est convexe par rapport à (z_1, z_2) pour tout $x \in \mathbb{R}$.

Pour $\xi = (\xi_1, \xi_2) \sim \mathcal{N}(0, I_2)$ on a que $g(\bar{x} := 0, \mu = 0) < 0$

φ est continu (grâce au point de Slater) mais **n'est même pas localement Lipschitz**



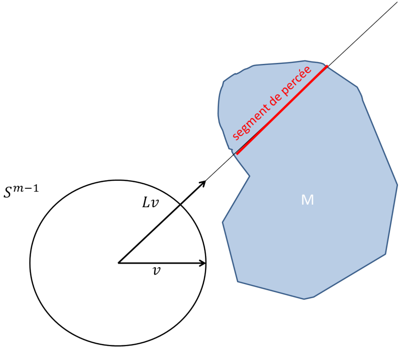
$$\varphi(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \int e^{-s^2/2} \Phi(1) ds & x \le 0 \\ \frac{1}{\sqrt{2\pi}} \int e^{-s^2/2} \Phi(1 - x^2 \exp(-1 - 4\log(1 - \Phi(s)))) ds & x > 0 \end{cases}$$

.

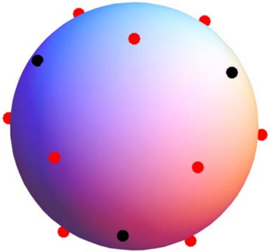
Soit $\xi \sim \mathcal{N}(0, R)$ avec $R = LL^T$. Alors,

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

où μ_η, μ_ζ sont les mesures de $\eta \sim \chi(m)$ et de la loi uniforme sur $\mathbb{S}^{m-1}$.



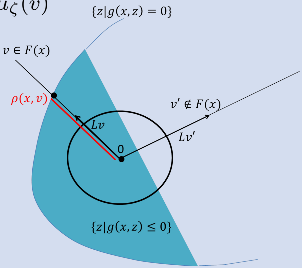
échantillonage de la sphère (e.g., QMC).



Theorem (v. Ackooij / H. 2014)

Supposons que $g \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R})$, $g(x,\cdot)$ *convexe,* $g(\bar{x},0) < 0$. *Alors,*

$$\varphi(\bar{x}) = \int_{v \in \mathbb{S}^{m-1}} \mu_\eta([0,\rho(\bar{x},v)]d\mu_\zeta(v)$$



Si en plus la condition de croissance modérée

$$\|\nabla_x g(x,\xi)\| \le l \exp(\|\xi\|) \, \forall x \in \mathbb{B}_{1/l}(\bar{x}) \, \forall \xi : \|\xi\| \ge l.$$

est satisfaite pour un $l > 0$, *alors*

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

En utilisant le même échantillon de la sphère on peut calculer à la fois $\varphi(\bar{x})$ et $\nabla\varphi(\bar{x})$.

Definition

Pour $g : X \times \mathbb{R}^m \to \mathbb{R}$ localement Lipschitz et $L > 0$, on définit le L- cône des directions modérées en $\bar{x} \in \mathbb{R}^n$, comme

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

Ici (dérivée directionnelle partielle au sens de Clarke),

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

Definition

Un cône fermé $K \subseteq X^*$ est dit d'avoir un w^*-compact sole s'il existe $x_0 \in X$ tel que

- $\langle x^*, x^0 \rangle > 0 \quad \forall x^* \in K \setminus \{0\}$
- l'ensemble $\{x^* \in K \mid \langle x^*, x^0 \rangle = 1\}$ est w^*- compacte.

Si $\dim X < \infty$, alors pour tout cône fermé $K \subseteq X$ on a que

$$\operatorname{int} K \neq \emptyset \iff K^* \text{ a un } w^*\text{- compact sole}$$

Theorem (Hantoute, H., Perez-Aros (2016))

Supposons que $\xi \sim \mathcal{N}(\mu, \Sigma)$ et que $g : X \times \mathbb{R}^m \to \mathbb{R}$ soit localement Lipschitz et convexe dans la deuxième variable. On fixe un point $\bar{x}$ avec $g(\bar{x}, \mu) < 0$. En plus, supposons que pour un $L > 0$ le dual C_L^* du L-cône des directions modérées en $\bar{x}$ possède un w^*-compact sole. Alors,

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

Ici, μ_ζ est la distribution uniforme sur $\mathbb{S}^{m-1}$ et

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

est la fonction probabiliste radiale ($\mu_\chi =$ loi χ avec m degrés de liberté).

Example

Dans notre contre-exemple nondifférentiable, on a que (pour $L > 0$ suffisamment large)

$$\partial\varphi(\bar{x}) = \{0\},\ C_L = (-\infty, 0],\ \partial_x e(\bar{x}, v) \subseteq (-\infty, 0]\ \forall v \in \mathbb{S}^{m-1},$$

d'où l'inclusion du Théorème se réduit ici à: $\{0\} \subseteq (-\infty, 0]$.

Theorem

Supposons que $\xi \sim \mathcal{N}(\mu, \Sigma)$ et que $g : X \times \mathbb{R}^m \to \mathbb{R}$ soit localement Lipschitz et convexe dans la deuxième variable. On fixe un point $\bar{x}$ avec $g(\bar{x}, \mu) < 0$. Si en plus $C_L = X$ pour un certain $L > 0$ où si l'ensemble $\{z \mid g(\bar{x}, z) \leq 0\}$ est borné, alors φ est localement Lipschitz autour de $\bar{x}$ et

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

Pour des fonctions localement Lipschitz f on a toujours que $\emptyset \neq \partial f(\bar{x})$ et

$$\#\partial f(\bar{x}) = 1 \iff f \text{ strictement différentiable à } \bar{x}$$

Corollary

Si on ajoute aux conditions du Théorème précédent que $\#\partial_x^C e(\bar{x}, v) = 1$ pour μ_ζ-a.e. v, alors φ est strictement différentiable à $\bar{x}$ et

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

Sousdifférentiel partiel de Clarke de la fonctionprobabiliste radiale $e(x,v)$

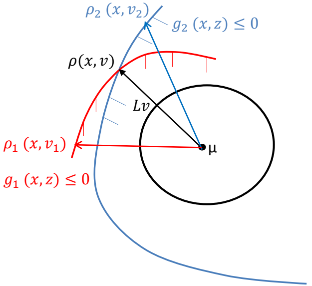
Theorem (v. Ackooij / H. 2015)

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

- $g_i \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^m)$ *et convexe dans la seconde variable*
- $g_i(\bar{x},\mu) < 0$ *pour* $i = 1,\ldots,p$ *(Slater point)*
- $C_L = X$ *pour un certain* $L > 0$.

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

Ici, $I(v) := \{i \mid \rho(\bar{x},v) = \rho_i(\bar{x},v)\}$ *et* χ *est la densité de la loi Chi avec* m *degrés de liberté.*



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

φ est strictement différentiable à $\bar{x}$.

Corollary

Si on ajoute au Théorème précédent la condition de qualification (constraint qualification) suivante

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

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

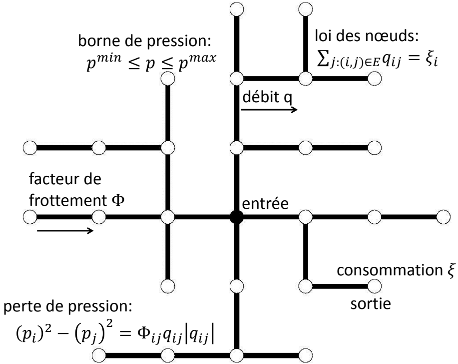
Alors, φ *est strictement différentiable à* $\bar{x}$*. Si cette condition reste vraie dans un voisinage de* $\bar{x}$*, alors* φ *est continûment différentiable. En plus, on a la formule du gradient*

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

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

Admissibilité d'une consommation aléatoire dans un réseau de gaz

Considérons un modèle très simple (passif, stationnaire) d'un réseau de gaz (V, E):



vecteur de consommation ξ admissible

$\Longleftrightarrow$

$\exists p, q$:

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

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

($A =$ matrice d'incidence)

Système d'inégalités explicite pour un réseau arborescent: vecteur de consommation ξ admissible $\Longleftrightarrow$

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

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

Contrainte mixte en probabilité et robustesse

Le propriétaire du réseau s'intéresse à garantir l'admissibilité de la consommation aléatoire avec une probabilité donnée:

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

Facteur de frottement Φ incertain aussi. A la différence de ξ on ne dispose pas d'une information statistique en générale. Modélisation du type 'pire des cas' par rapport à un ensemble d'incertitude:

$$\mathbb{P}\left((p_k^{\max})^2 + g_k(\xi,\Phi) \geq (p_l^{\min})^2 + g_l(\xi,\Phi) \quad (k,l = 0,\ldots,|V|),\ \forall\Phi \in [\bar{\Phi}-\delta,\bar{\Phi}+\delta]\right) \geq p \tag{1}$$

Ici, $\bar{\Phi}$ est un vecteur nominal de facteurs de frottement.

Système infini d'inégalités aléatoires. Modèle mixte de contrainte en proba et de contrainte robuste.

Choix de δ souvent pas évident. Pour gagner d'information sur la sensibilité locale de l'incertitude en Φ:

'Maximiser' l'ensemble d'incertitude tout en gardant l'admissibilité de la consommation avec une probabilité donnée:

$$\text{maximiser} \quad \sum_{e\in E} \delta_e^{0.9} \quad \text{sous contrainte en proba (1)}$$

Résolution numérique pour un exemple

Illustration de la solution optimale pour un arbre avec 27 noeuds, $p = 0.8$, ξ Gaussien: