Conception de détecteurs quantiques optimaux via le calcul par intervalles

Journée du GdR CNRS ISIS - Traitement du signal et applications quantiques
22 juin 2021

Pierre Engelstein, Nicolas Delanoue, François Chapeau-Blondeau

Sommaire

Opérateur densité

Définition

Un *opérateur densité* pour un état quantique est une matrice Hermitiene ρ, telle que $\rho = \rho^\dagger$, de trace $\mathrm{tr}(\rho) = 1$ et semi-définie positive $\rho \succeq 0$.

Définition

Soit un système quantique dans un état **pur** $|\phi_j\rangle$. L'opérateur densité correspondant à cet état est définit par :

$$\rho_j = |\psi_j\rangle\langle\psi_j| \,. \tag{1}$$

L'opérateur densité est aussi capable de représenter des états **mixtes** qui ne sont pas représentables par les vecteurs d'états.

Exemple

Soit un état quantique

$$|\psi_j\rangle = |+\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

L'opérateur densité correspondant est :

$$\rho_j = |\psi_j\rangle\langle\psi_j| = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Opérateurs de mesure [1]

Définition

La *mesure projective* est définie par un ensemble de N projecteurs orthogonaux $|n\rangle\langle n| = \Pi_n$, vérifiant $\sum_n |n\rangle\langle n| = \Pi_n = I_N$ et $Pr\{|n\rangle\} = \mathrm{tr}(\rho\Pi_n)$.

Définition : Mesure généralisée (POVM)

On généralise la notion de mesure projective à un ensemble non orthogonal d'opérateurs de mesure appelés POVM (*positive operator-valued measurement*) $\{\Pi_n\}$ vérifiant $\sum_n \Pi_n = I_N$ et $Pr\{\Pi_n\} = \mathrm{tr}(\rho\Pi_n)$.

1. Chapeau-Blondeau, F., "Quantum information, quantum computation : An introduction", Cours de formation doctorale (2018)

Matrice de probabilités [2]

Définition

Soit un ensemble d'états quantiques $\{\rho_i\}$ avec leur probabilités $\{p_i\}$ tel que $\sum_i p_i = 1$.
Considérons un ensemble correspondant de POVM $\{\Pi_j\}$ permettant de réaliser une mesure généralisée.
La matrice de probabilités est définie par :

$$P_{ij} = p_i \operatorname{tr}(\rho_i \Pi_j) \tag{2}$$

Interprétation

Un élément P_{ij} correspond à la probabilité de mesurer Π_j en ayant l'état quantique ρ_i.

2. Davies, E., "Information and quantum measurement", IEEE Transaction on Information Theory 24 (1978), 596-599

Critères d'optimisation

1. Minimiser une combinaison linéaire des probabilités (cf. refs 6, 7, 9, 10, 15 davies)
2. Maximiser la probabilité de détection correcte (linéaire), [3]
3. Minimiser l'erreur quadratique de mesure, [4]
4. Maximiser information mutuelle [5].

---

3. Eldar, Y., "Designing Optimal Quantum Detectors Via Semidefinite Programming", IEEE Transaction on Information Theory 49 (2003), 1012-1017

4. Eldar, Y., "On Quantum Detection and the Square-Root Measurement", IEEE Transaction on Information Theory 47 (2001), 858-872

5. Davies, E., "Information and quantum measurement", IEEE Transaction on Information Theory 24 (1978), 596-599

Maximiser l'information mutuelle

L'*information mutuelle* est définie pour deux variables aléatoires X et Y par :

$$\begin{aligned} I(X;Y) &= H(X) + H(Y) - H(X,Y) && (3)\\ &= H(Y) - H(Y|X) && (4)\\ &= H(X) - H(X|Y) && (5)\end{aligned}$$

1. $H(X)$ et $H(Y)$ entropies marginales de X et Y ;
2. $H(X,Y)$ entropie conjointe de X et Y ;
3. $H(X|Y)$ entropie de X sachant Y.

Définition

L'*information mutuelle* est définie par rapport à la matrice des probabilités P :

$$I(P) = \sum_i H(\sum_j P_{ij}) + \sum_j H(\sum_i P_{ij}) - \sum_{ij} H(P_{ij}) \qquad (6)$$

Avec $H(x) = -x\log(x)$ entropie de la variable x.

Maximiser l'information mutuelle

Problème

On cherche à résoudre le problème de maximisation :

$$\max_{\Pi} I(\rho, \Pi) \tag{3}$$

tel que :

$$\Pi_j \succeq 0, \quad 1 \leq j \leq m \tag{4}$$

$$\sum_{j=1}^{m} \Pi_j = I \tag{5}$$

- Problème de maximisation de fonction convexe : difficile en pratique à résoudre.
- Utilisation du calcul par intervalle pour fournir une solution optimale garantie.

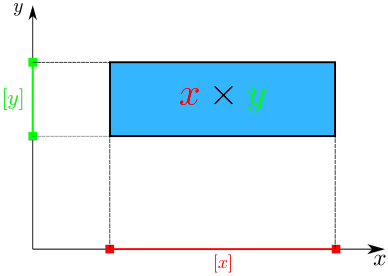
Intervalles

Notion d'intervalle

Un intervalle $\mathbf{x} = [\underline{x}, \overline{x}]$ est défini comme l'ensemble des nombres réels x t.q. $\underline{x} \leq x \leq \overline{x}$.

Notion de boites

Une boite $\mathbf{X} = (\mathbf{x1}, \mathbf{x2}, ..., \mathbf{xj})$ est le produit cartésien des intervalles $\mathbf{x1} \times \mathbf{x2} \times \cdots \times \mathbf{xj}$

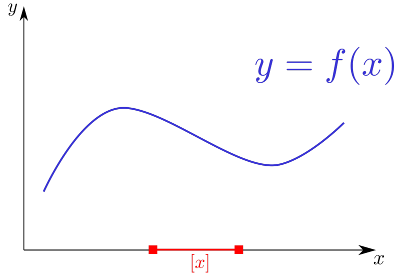


Fonction d'inclusion

Définition

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction, la fonction $[f] : \mathbb{R}^n \to \mathbb{R}^m$ est une *fonction d'inclusion* pour f si

$$\forall [\underline{x}, \overline{x}] \in \mathbb{R}^n, f([\underline{x}, \overline{x}]) \subset [f]([\underline{x}, \overline{x}]) \tag{6}$$

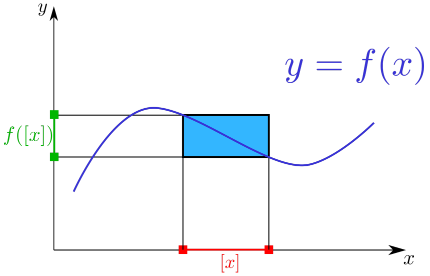


Fonction d'inclusion

Définition

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction, la fonction $[f] : \mathbb{R}^n \to \mathbb{R}^m$ est une *fonction d'inclusion* pour f si

$$\forall [\underline{x}, \overline{x}] \in \mathbb{R}^n, f([\underline{x}, \overline{x}]) \subset [f]([\underline{x}, \overline{x}]) \quad (6)$$

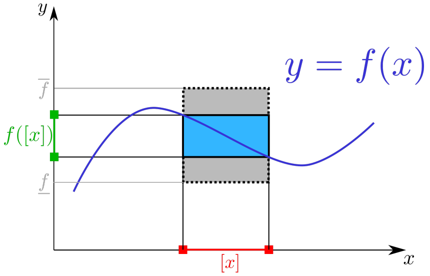


Fonction d'inclusion

Définition

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction, la fonction $[f] : \mathbb{R}^n \to \mathbb{R}^m$ est une *fonction d'inclusion* pour f si

$$\forall[\underline{x}, \overline{x}] \in \mathbb{R}^n, f([\underline{x}, \overline{x}]) \subset [f]([\underline{x}, \overline{x}]) \tag{6}$$



Arithmétique des intervalles

On définit un ensemble d'opérateurs comme extension des opérateurs arithmétiques sur les nombres réels :

- $[x_1] + [x_2] = [\underline{x_1} + \underline{x_2}, \overline{x_1} + \overline{x_2}]$
- $[x_1] - [x_2] = [\underline{x_1} - \overline{x_2}, \overline{x_1} - \underline{x_2}]$
- $[x_1] \times [x_2] = [\min(\underline{x_1 x_2}, \underline{x_1}\overline{x_2}, \overline{x_1}\underline{x_2}, \overline{x_1 x_2}), \max(\underline{x_1 x_2}, \underline{x_1}\overline{x_2}, \overline{x_1}\underline{x_2}, \overline{x_1 x_2})]$
- $\log([x]) = [\log(\underline{x_1}), \log(\overline{x_2})]$
- ...

Exemple

$$\begin{aligned} [z] &= ([2,3] - [-4,5])^2 + [-3,2] \\ &= [-3,7]^2 + [-3,2] \\ &= [0,49] + [-3,2] \\ &= [-3,51] \end{aligned}$$

Petite histoire du calcul par intervalle [6]

1. Arithmétique des intervalles, R. E. Moore, 1966,
2. Optimisation globale, R. Baker Kearfott, 90's,
3. Résolution de systèmes d'équations, A. Neumaier, 90's,
4. Résolution d'équations différentielles ordinaires, R. Lohner 1988,
5. La preuve de l'existence de l'attracteur étrange pour le système de Lorentz, W. Tucker, 1998,
6. Analyse par intervalles appliquée à la robotique, L. Jaulin, 2001,
7. La preuve de la conjecture de Kepler, T. Hales, 2003,
8. Estimation de paramètres pour les systèmes décrits par les équationsdifférentielles ordinaires, N. Ramdani, 2004,
9. EDP, topologie algébrique, . . .

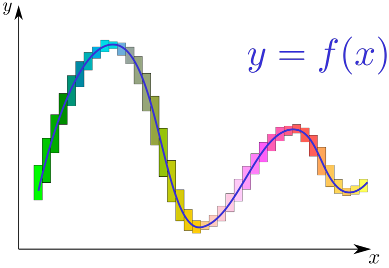
6. Delanoue, N., "Méthodes numériques garanties pour la classification de fonctions et le contrôle optimal", Soutenance HDR, 2018

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max_x f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \quad (7)$$



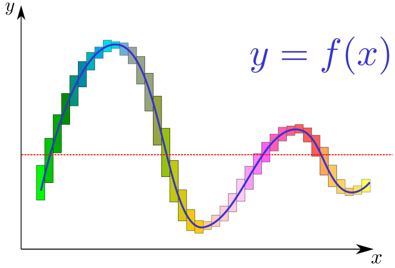
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max\limits_{x} f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \tag{7}$$



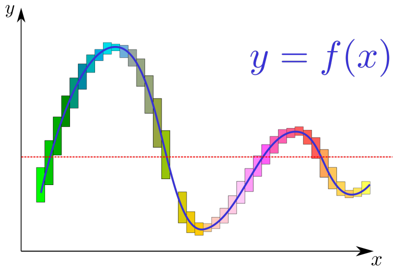
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max_x f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \tag{7}$$



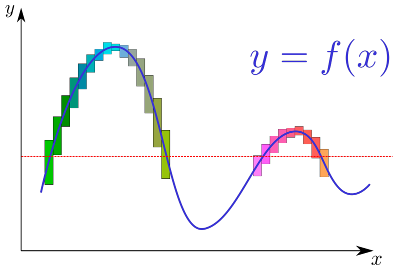
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max\limits_{x} f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \tag{7}$$



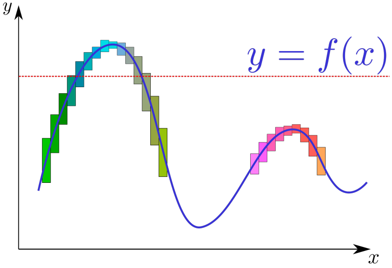
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max_x f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \quad (7)$$



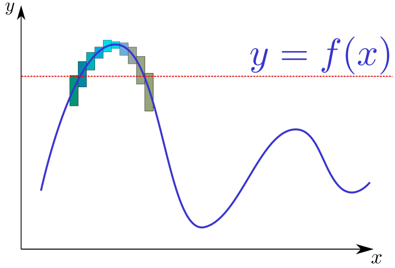
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max_x f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \tag{7}$$



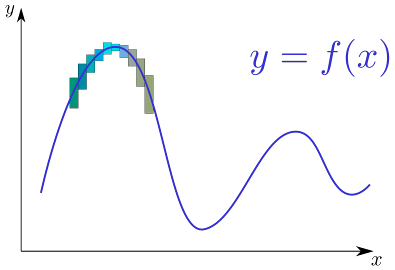
$y = f(x)$

Premier algorithme d'optimisation

Théorème

Soit un a une solution admissible du problème $\max_x f(x)$ tel que $g(x) \leq 0$ et x^* la solution optimale, on a :

$$sup([f]([x])) \leq f(a) \Rightarrow x^* \notin [x] \quad (7)$$



$y = f(x)$

Algorithme avancé

Quatres étapes pour l'algorithme, itérativement :

- Bisection d'une boite selon un axe donné (plus grand axe par exemple) ;
- Évaluation des contraintes sur les boites résultantes : permet d'éliminer les boites hors-contraintes ;
- Borner la fonction avec le calcul par intervalle ;
- Choix d'un candidat *a* pour l'élimination par borne supérieure (point milieu, recherche d'un maximum dans l'intervalle avec un algorithme classique, ...)

Exemple concret

Considérons deux états purs $|\psi_1\rangle = \begin{bmatrix} \frac{1}{3} \\ \frac{2\sqrt{(2)}}{3} \end{bmatrix}$ et $|\psi_2\rangle = |+\rangle$, avec $p_1 = 0.1$ et $p_2 = 0.9$.
Les opérateurs densité correspondant sont :

$$\rho_1 = \begin{bmatrix} \frac{1}{9} & \frac{2\sqrt{2}}{9} \\ \frac{2\sqrt{2}}{9} & \frac{8}{9} \end{bmatrix}, \quad \rho_2 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

Et on a deux opérateurs de mesure inconnus soit 8 variables inconnues :

$$\Pi_1 = \begin{bmatrix} a_1 & b_1 + ic_1 \\ b_1 - ic_1 & d_1 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} a_2 & b_2 + ic_2 \\ b_2 - ic_2 & d_2 \end{bmatrix}$$

On veut résoudre le problème

$$\max_{\Pi_1, \Pi_2} I(\rho_1, \rho_2, \Pi_1, \Pi_2) \tag{8}$$

tel que :

$$\Pi_1 \succeq 0, \Pi_2 \succeq 0 \tag{9}$$
$$\Pi_1 + \Pi_2 = I \tag{10}$$

Exemple concret

On sait que $\Pi_1 + \Pi_2 = I_2$, on réduit à 4 variables inconnues :

$$\Pi_2 = \begin{bmatrix} (1 - a_1) & (-b_1) + i(-c_1) \\ (-b_1) - i(-c_1) & (1 - d_1) \end{bmatrix}$$

Comme les ρ_i ne comportent pas de termes complexes, on peut retirer le terme c des Π_j, on réduit à 3 variables $\{a, b, d\}$ notre problème.
On pose les contraintes de Silvester pour la contrainte de semi-définition positive des opérateurs de mesure :

$$\begin{aligned} a &\geq 0 \\ d &\geq 0 \\ a \times d - b^2 &\geq 0 \\ (1 - a) \times (1 - d) - (-b)^2 &\geq 0 \end{aligned}$$

Résultats

Ibex

Ibex nous donne $I = 0.472\dots$ avec une précision de 4×10^{-5}, et
$\Pi_1 = \begin{bmatrix} 0.454 & -0.498 \\ -0.498 & 0.546 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.546 & 0.498 \\ 0.498 & 0.454 \end{bmatrix}.$
Le temps de calcul est de 88 secondes.

Notre solveur

Notre solveur nous donne $I = 0.472\dots$ avec une précision de 1×10^{-9}, et
$\Pi_1 = \begin{bmatrix} 0.456 & -0.498 \\ -0.498 & 0.544 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.544 & 0.498 \\ 0.498 & 0.456 \end{bmatrix}.$
Le temps de calcul est de 10 secondes.

Perspectives

1. Étendre le problème à des états d'entrée bruités : on considère alors les $\{\rho\}$ comme étant inconnus ;
2. Difficulté d'étendre le problème à des dimensions supérieures : complexité exponentielle avec la dimension.
3. L'algorithme étant parallélisable, on peut néanmoins réduire considérablement le temps de calcul, en fonction du nombre de processeurs.