MATEMÁTICA UNIVERSITÁRIA
n$^{\underline{o}}$32 – junho 2002 – pp. 1–7

A Projeção e seu Potencial

Rolci Cipolatti

É sabido que o operador de projeção sobre um convexo fechado de $\mathbb{R}^n$ (ou mais geralmente, um convexo fechado de um espaço de Hilbert H) desempenha um papel importante na Matemática e nas suas aplicações. É dele que queremos falar e dirigimo-nos em especial aos estudantes.

Se C é um conjunto convexo, fechado e não vazio de $\mathbb{R}^n$, então, como precisaremos adiante, podemos definir para cada $x \in \mathbb{R}^n$ a sua projeção sobre C, que será denotada por $P_C(x)$. Temos, assim, definida a função $P_C\colon \mathbb{R}^n \to \mathbb{R}^n$ — o Operador Projeção sobre C.

O objetivo desta nota é mostrar que podemos construir explicitamente uma função $J\colon \mathbb{R}^n \to \mathbb{R}$ — o Potencial da Projeção — que é convexa, de classe C^1 e tal que

$$J'(x) = P_C(x), \quad \forall x \in \mathbb{R}^n.$$

Mais precisamente, se $\langle x; y\rangle$ denota o produto escalar usual de $\mathbb{R}^n$, temos:

Teorema 1 *Seja C um conjunto convexo, fechado e não vazio de $\mathbb{R}^n$ e considere a função $J\colon \mathbb{R}^n \to \mathbb{R}$ definida por*

$$J(x) := \big\langle x - \frac{1}{2}P_C(x); P_C(x)\big\rangle, \tag{1}$$

onde $P_C\colon \mathbb{R}^n \to \mathbb{R}^n$ é a projeção sobre C. Então J é função convexa de classe C^1 em $\mathbb{R}^n$ e $J' = P_C$.

Antes de passarmos à prova do Teorema 1, sejamos mais precisos. A proposição a seguir define a função projeção e aponta algumas de suas propriedades mais importantes.

Se $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, denotamos

$$\|x\| = \sqrt{\langle x; x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Proposição 1 *Seja $C \subset \mathbb{R}^n$ um conjunto convexo, fechado e não vazio. Então:*

(a) *Para todo $x \in \mathbb{R}^n$, existe um único $y \in C$ tal que*

$$\|x - y\| \leq \|z - x\|, \quad \forall z \in C. \tag{2}$$

$y = P_C(x)$ *é denominado a* Projeção *de x sobre C. Temos assim definida a aplicação*

$$\begin{array}{rcl} P_C\colon \mathbb{R}^n & \to & \mathbb{R}^n \\ x & \mapsto & P_C(x) \end{array} \tag{3}$$

(b) $y = P_C(x) \iff \langle x - y\,;\, z - y\rangle \leq 0,\ \forall z \in C.$

(c) *Para todo $x_1, x_2 \in \mathbb{R}^n$, temos*

$$\|P_C(x_1) - P_C(x_2)\|^2 \leq \big\langle P_C(x_1) - P_C(x_2); x_1 - x_2\big\rangle. \tag{4}$$

Em particular, P_C é função Lipschitz-contínua em $\mathbb{R}^n$.

Prova: Seja $x \in \mathbb{R}^n$. Se $x \in C$, então $y = x$ satisfaz (2). Se $x \notin C$, seja $x_1 \in C$ e considere $r = \|x - x_1\| > 0$. É claro que $C_r := \overline{B_r(x)} \cap C$ é compacto e não vazio. Como a função $z \mapsto \|x - z\|$ é contínua, existe $y \in C_r$ tal que $\|x-y\| \leq \|x-z\|$, $\forall z \in C_r$. Por outro lado, se $z \in C \setminus C_r$, então

$$\|x - z\| \geq \|x - x_1\| \geq \|x - y\|$$

e obtemos a desigualdade (2).

Seja $z \in C$. Então, para todo $t \in]0, 1[$ temos $(1-t)y + tz \in C$ e, em particular, $\|x - y\|^2 \leq \|x - (1-t)y - tz\|^2$, o que implica

$$\langle x - y; z - y\rangle \leq \frac{t}{2}\|y - z\|^2.$$

Fazendo $t \to 0^+$, obtemos a desigualdade em (b).

Para provar que y é único, suponhamos y_1, y_2 satisfazendo (2). Então

$$\begin{aligned} \langle x - y_1; z - y_1\rangle &\leq 0 \\ \langle x - y_2; z - y_2\rangle &\leq 0 \end{aligned} \quad \forall z \in C. \tag{5}$$

Substituindo $z = y_2$ na primeira desigualdade de (5), $z = y_1$ na segunda e somando as duas, obtemos $\|y_1 - y_2\| \leq 0$, ou $y_1 = y_2$.

Para provar (c), consideremos $x_1, x_2 \in \mathbb{R}^n$. Então segue de (b) que

$$\begin{aligned} \langle x_1 - P_C(x_1); P_C(x_2) - P_C(x_1)\rangle &\leq 0, \\ \langle x_2 - P_C(x_2); P_C(x_1) - P_C(x_2)\rangle &\leq 0. \end{aligned} \tag{6}$$

Somando as desigualdade em (6), obtemos (c).

Para mostrar que P_C é Lipschitz-contínua, basta aplicar a desigualdade de Cauchy-Schwarz no lado direito de (4). □

Sabemos do Cálculo Diferencial (veja [3]) que se $f\colon \mathbb{R} \to \mathbb{R}$ é derivável, então f é convexa se e somente se f' é função monótona crescente. Esta propriedade pode ser generalizada para funções $f\colon \mathbb{R}^n \to \mathbb{R}$ se considerarmos a extensão apropriada do conceito de *função crescente* para funções vetoriais.

Definição 1 *Uma função $g\colon \mathbb{R}^n \to \mathbb{R}^n$ é dita monótona positiva em $\mathbb{R}^n$ se*

$$\big\langle g(x) - g(y); x - y\big\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^n.$$

O proposição seguinte (veja [2]) nos será útil para a prova do Teorema 1.

Proposição 2 *Seja $f\colon \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável. Então f é convexa se e somente se $g = f'$ é monótona positiva (onde $f'(x_0)$ denota o gradiente de f em x_0).*

Prova: Provemos inicialmente a implicação "⇒". Por hipótese temos, para $t \in [0, 1]$,

$$f\big(x_0 + t(x_1 - x_0)\big) \leq f(x_0) + t\big(f(x_1) - f(x_0)\big),$$
$$f\big(x_0 + t(x_1 - x_0)\big) = f(x_0) + t\big\langle f'(x_0); x_1 - x_0\big\rangle + \epsilon\big(t(x_1 - x_0)\big),$$

onde
$$\lim_{\xi\to 0}\frac{\epsilon(\xi)}{\|\xi\|}=0,\ \xi\in\mathbb{R}^n.$$
Subtraindo a segunda equação da primeira, obtemos
$$t\big(f(x_1)-f(x_0)\big)\geq t\big\langle f'(x_0);x_1-x_0\big\rangle+\epsilon\big(t(x_1-x_0)\big).$$
Denotando por $\xi=t(x_1-x_0)$, $t>0$, temos após divisão por t
$$f(x_1)-f(x_0)\geq\big\langle f'(x_0);x_1-x_0\big\rangle+\frac{\epsilon(\xi)}{\|\xi\|}\|x_1-x_0\|.$$
Fazendo $t\to 0$, concluímos
$$f(x_1)-f(x_0)\geq\big\langle f'(x_0);x_1-x_0\big\rangle.$$
Mutatis mutandis,
$$f(x_0)-f(x_1)\geq\big\langle f'(x_1);x_0-x_1\big\rangle$$
e temos a conclusão.

Provemos a implicação contrária "$\Leftarrow$". Sabemos da Análise Real que se $\varphi\colon\mathbb{R}\to\mathbb{R}$ é derivável e φ' é crescente, então φ é convexa. Sejam $x_1,x_0\in\mathbb{R}^n$ e consideremos $\varphi(t)=f\big(x_0+t(x_1-x_0)\big)$. Como f é diferenciável, segue da Regra da Cadeia que $\varphi'(t)=\big\langle f'\big(x_0+t(x_1-x_0)\big);x_1-x_0\big\rangle$. Provemos que φ' é crescente.
$$\varphi'(t_1)-\varphi'(t_0)=\big\langle f'\big(x_0+t_0(x_1-x_0)\big)-f'\big(x_0+t_1(x_1-x_0)\big);x_1-x_0\big\rangle.$$
Como $\big(x_0+t_1(x_1-x_0)\big)-\big(x_0+t_0(x_1-x_0)\big)=(t_1-t_0)(x_1-x_0)$, podemos escrever
$$(t_1-t_0)\big(\varphi'(t_1)-\varphi'(t_0)\big)=\big\langle f'(x_{t_1})-f'(x_{t_0});x_{t_1}-x_{t_0}\big\rangle,$$
onde estamos denotando $x_t=x_0+t(x_1-x_0)$.

Como por hipótese f' é monótona positiva, concluímos que φ' é crescente. Logo φ é convexa e $\varphi(t)\leq\varphi(0)+t(\varphi(1)-\varphi(0))$ para todo $t\in\,]0,1[$. Portanto,
$$f\big(x_0+t(x_1-x_0)\big)\leq f(x_0)+t\big(f(x_1)-f(x_0)\big)$$

para todo $t \in]0,1[$. □

Prova do Teorema 1: Sejam x_0 e h em $\mathbb{R}^n$. Então podemos escrever

$$J(x_0+h) = J(x_0) + \langle P_C(x_0); h \rangle + \varepsilon(h),$$

onde

$$\varepsilon(h) := \frac{1}{2}\|P_C(x_0)\|^2 - \frac{1}{2}\|P_C(x_0+h)\|^2 + \langle x_0+h; P_C(x_0+h) - P_C(x_0) \rangle.$$

Como

$$\|P_C(x_0)\|^2 - \|P_C(x_0+h)\|^2 = \langle P_C(x_0)+P_C(x_0+h); P_C(x_0)-P_C(x_0+h) \rangle,$$

podemos escrever

$$\begin{aligned} \varepsilon(h) &= \langle x_0 - P_C(x_0); P_C(x_0+h) - P_C(x_0) \rangle \\ &\quad -\tfrac{1}{2}\|P_C(x_0+h) - P_C(x_0)\|^2 + \langle h; P_C(x_0+h) - P_C(x_0) \rangle. \end{aligned} \tag{7}$$

Considerando $x_1 = x_0$ e $x_2 = x_0 + h$ na primeira desigualdade de (6), obtemos

$$\langle x_0 - P_C(x_0); P_C(x_0+h) - P_C(x_0) \rangle \leq 0.$$

Logo

$$\varepsilon(h) \leq \langle h; P_C(x_0+h) - P_C(x_0) \rangle.$$

Aplicando a desigualdade de Cauchy-Schwarz e o item (c) da Proposição 1, obtemos $\varepsilon(h) \leq \|h\|^2$.

Por outro lado, podemos escrever também que

$$\begin{aligned} \varepsilon(h) &= \langle x_0 + h - P_C(x_0+h); P_C(x_0+h) - P_C(x_0) \rangle \\ &\quad +\tfrac{1}{2}\|P_C(x_0) - P_C(x_0+h)\|^2. \end{aligned}$$

Usando a segunda desigualdade de (6) (com os mesmos x_1 e x_2 definidos acima), obtemos $\varepsilon(h) \geq 0$.

Portanto, $0 \leq \varepsilon(h) \leq \|h\|^2$, $\forall h \in \mathbb{R}^n$ e concluímos que

$$J'(x_0) = P_C(x_0). \tag{8}$$

A prova de que J é de classe C^1 decorre diretamente de (8) e do item (c) da Proposição 1.

A prova de que J é convexa decorre diretamente da Proposição 2 e do item (c) da Proposição 1. □

Corolário 1 *Seja C um conjunto convexo, fechado e não vazio de $\mathbb{R}^n$ e considere a função $G\colon \mathbb{R}^n \to \mathbb{R}$ definida por*

$$G(x) := \frac{1}{2}\|x - P_C(x)\|^2.$$

Então G é função convexa de classe C^1 em $\mathbb{R}^n$ e $G'(x) = x - P_C(x)$.

Prova: Como $G(x) = \frac{1}{2}\|x\|^2 - J(x)$ e a função $f(x) := \frac{1}{2}\|x\|^2$ é de classe C^1 com $f'(x) = x$ para todo $x \in \mathbb{R}^n$, segue do Teorema 1 que G é de classe C^1 e $G'(x) = x - P_C(x)$. Além disso, segue de (4) que

$$\langle G'(x_1) - G'(x_2); x_1 - x_2 \rangle \geq 0.$$

Portanto, G' é monótona positiva e concluímos que G é convexa. □

Observação: Como $P_C(x)$ é o ponto de C mais próximo de x, a função

$$D(x) := \|x - P_C(x)\| = \sqrt{2G(x)}$$

mede a distância de x a C. Segue da regra da cadeia e do Corolário 1 que a função D é diferenciável no complementar de C e, para todo todo $x \in \mathbb{R}^n \setminus C$,

$$D'(x) = \frac{x - P_C(x)}{\|x - P_C(x)\|}$$

é o vetor unitário na direção de maior crescimento dessa distância.

Exemplo 1 *Seja $C = [0, 1]$. Então*

$$P_C(x) = \begin{cases} 0 & se\ x < 0 \\ x & se\ 0 \leq x \leq 1 \\ 1 & se\ x > 0 \end{cases} \quad \text{e} \quad J(x) = \begin{cases} 0 & se\ x < 0 \\ x^2/2 & se\ 0 \leq x \leq 1 \\ x - 1/2 & se\ x > 0 \end{cases}$$

Exemplo 2 *Seja $C = \overline{B_1(0)}$ a bola fechada de raio 1 de $\mathbb{R}^n$. Então*

$$P_C(x) = \begin{cases} x & se\ \|x\| \leq 1 \\ x/\|x\| & se\ \|x\| > 1 \end{cases}$$

e

$$J(x) = \begin{cases} \|x\|^2/2 & se\ \|x\| \leq 1 \\ \|x\| - 1/2 & se\ \|x\| > 1 \end{cases}$$

Observação: Embora estejamos nos restringindo a $\mathbb{R}^n$ munido do produto escalar usual, podemos constatar que os resultados valem com provas idênticas para qualquer espaço de Hilbert H, exceto a prova da existência de y no item (a) da Proposição 1, onde usamos argumentos de compacidade que não valem se H é de dimensão infinita. Nesse caso, a prova da existência de y satisfazendo (2) pode ser obtida usando-se resultados básicos da Análise Funcional (veja [1], Teorema V.2, p. 79).

Referências

[1] H. Brezis, Analyse Fonctionnelle, Théorie et Applications, Collection Mathématiques Appliquées pour la Maîtrise, Masson, 1983.

[2] R. Cipolatti, Cálculo Avançado I, Edit. Instituto de Matemática - UFRJ, Rio de Janeiro, 2001.

[3] M. Spivak, Calculus, W. A. Benjamin, Inc., New York, 1970.

Departamento de Métodos Matemáticos
Instituto de Matemática, Universidade Federal do Rio de Janeiro
C.P. 68530, Rio de Janeiro, Brasil
e-mail: cipolatti@im.ufrj.br