Nota sobre o Melhor Par Aproximante de Duas Variedades Lineares Enviesadas

M. A. Facas Vicente, *José Vitória*
Departamento de Matemática, Universidade de Coimbra
e-mail: vicente@mat.uc.pt

P. Saraiva
CMUC & CeBER, Universidade de Coimbra
e-mail: psaraiva@fe.uc.pt

P. D. Beites[1]
CMA & DM, Universidade da Beira Interior
e-mail: pbeites@ubi.pt

Armando Gonçalves
IPC, ESEC, MEM, Coimbra
e-mail: adsgoncalves@esec.pt

Resumo: Nesta nota obtém-se o melhor par aproximante de duas variedades lineares enviesadas após determinar a projeção de um ponto sobre uma variedade linear. Concretamente, calculam-se por duas vezes os vetores de norma mínima, adequados em cada caso, pertencentes a duas variedades lineares paralelas. Obtêm-se ainda expressões para as referidas projeções envolvendo a inversa de Moore-Penrose de cada uma das matrizes associadas às variedades lineares consideradas (na sua formulação matricial).

Abstract: The best approximation pair of two skew linear varieties is obtained after getting the projection of a point onto a linear variety. Concretely, we compute twice the minimum norm vectors of two, adequate each time, parallel linear varieties. In addition, we obtain expressions for the mentioned projections by means of the Moore-Penrose inverse of each of the matrices associated with the linear varieties considered (in its matrix version).

palavras-chave: variedades lineares enviesadas; melhor par aproximante; teoria da aproximação; vetor de norma mínima; inversa de Moore-Penrose.

keywords: skew linear varieties; best approximation pair; approximation theory; minimum norm vector; Moore-Penrose inverse.

[1] Financiada por Fundação para a Ciência e a Tecnologia (Portugal), projeto UID/MAT/00212/2013, e Ministerio de Economía y Competitividad (España), projeto MTM2013-45588-C3-1-P.

1 Introdução

A presente nota aborda dois problemas clássicos de Geometria Analítica Euclidiana n-dimensional – determinar a projeção ortogonal de um ponto sobre uma dada variedade linear e obter uma expressão para a distância entre duas variedades lineares enviesadas – de um modo que justifica revisitá-los. Com efeito, no primeiro resultado principal desta nota, recorremos aos vetores de norma mínima pertencentes a duas variedades lineares paralelas. O segundo problema principal diz respeito à obtenção do melhor par aproximante de pontos pertencentes a duas variedades lineares enviesadas, mediante a resolução de um sistema de equações lineares, construído por aplicação do primeiro resultado. Ao obtermos o melhor par aproximante, vamos mais longe do que o habitual, pois em [2], [4] e [7] apenas se focou o problema da distância entre duas variedade lineares. De facto, a julgar por estas referências, os pontos que materializam tal distância não parecem ter sido considerados.

O problema da determinação da (mínima) distância euclidiana entre duas variedades lineares foi também recentemente tratado em [6] e [5], aplicando, respetivamente, a Teoria de Gram e a inversa de Moore-Penrose de uma matriz particionada. Ainda que nestas referências se exiba o melhor par aproximante de pontos, as abordagens aí utilizadas são distintas da que se aplica na presente nota. Além disso, os resultados agora apresentados situam-se no contexto dos problemas de projeções sobre certos conjuntos convexos, baseados portanto em resultados sobre existência, unicidade e caracterização de soluções para problemas de melhor aproximação. Também no presente trabalho se recorre à inversa de Moore-Penrose, a qual constitui um excelente instrumento teórico (veja-se, *e.g.*, [1] e [11]). Contudo, é justo referir que, do ponto de vista numérico, tal ferramenta contém algumas fragilidades: a inversa de Moore-Penrose não é contínua e não é computacionalmente estável [10, pp. 423–424].

O presente trabalho está organizado do seguinte modo. Começamos por enunciar algumas definições, notações e resultados, coligidos na secção 2. Na secção 3, apresentamos os resultados sobre a projeção de um ponto sobre uma variedade linear. Na secção 4, recorrendo aos resultados da anterior secção, apresentamos um método original para obter o melhor par aproximante de duas variedades lineares enviesadas.

2 Preliminares

Ao longo da presente nota, $\mathbb{R}^n$ designa o espaço vetorial real n-dimensional euclidiano usual. O *produto interno* será denotado por $\bullet$ e define-se, como habitualmente, do seguinte modo: dados $\vec{p} = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix}^T$ e $\vec{q} = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}^T \in \mathbb{R}^{n\times 1}$, onde T designa a transposta, tem-se

$$\vec{p} \bullet \vec{q} = \sum_{i=1}^{n} p_i q_i.$$

Recorde-se também que a *norma euclidiana* de $\vec{p}$ é dada por

$$\|\vec{p}\| = \sqrt{\vec{p} \bullet \vec{p}}.$$

Além disso, considerando o espaço afim associado a $\mathbb{R}^n$, a extremidade de cada vetor posicional (relativamente à origem das coordenadas) será identificada com o próprio vetor.

Dados um ponto $\vec{a}$ de $\mathbb{R}^n$ e um subespaço N de $\mathbb{R}^n$, de dimensão m, o conjunto

$$V_{\vec{a}} = \vec{a} + N$$

diz-se uma *variedade linear*, [9]. Uma variedade linear pode então ser entendida como o resultado da translação de um subespaço.

De entre as várias maneiras de representar uma variedade linear (ver [3], [8] e [9]), recorreremos ainda à *versão matricial*, a qual possibilita a utilização do método cartesiano. Assim, dada uma variedade linear $V_{\vec{a}} = \vec{a} + N$, para alguma matriz $A \in \mathbb{R}^{p\times n}$, com característica completa por linhas, e para algum $\vec{c} \in \mathbb{R}^p$ com $p \in \{1, \ldots, n\}$, tem-se

$$V_{\vec{a}} = \{\vec{x} \in \mathbb{R}^n : \ A\vec{x} = \vec{c}\}.$$

Relativamente à representação dada na definição, é importante reter que

$$N = \mathcal{N}(A) \text{ (i.e., o núcleo de } A)$$

e que p é a codimensão de N ($p \leq n$). Tendo isto em conta, doravante iremos identificar uma variedade linear $\mathcal{L}$ dos seguintes dois modos:

$$\mathcal{L} := \{\vec{x} \in \mathbb{R}^n : L\vec{x} = \vec{c}\} = \vec{l}_0 + \mathcal{N}(L),$$

apenas optando por uma delas se a outra se verificar desnecessária.

Sejam dados pontos $\vec{l}_0, \vec{m}_0 \in \mathbb{R}^n$ e consideremos duas variedade lineares

$$\mathcal{L} = \ \vec{l}_0 + \mathcal{N}(L) \text{ e } \mathcal{M} = \ \vec{m}_0 + \mathcal{N}(M).$$

Seguindo [4], as variedades $\mathcal{L}$ e $\mathcal{M}$ dizem-se *paralelas* se $\mathcal{N}(L) = \mathcal{N}(M)$.

Por outro lado, admitamos que as variedades $\mathcal{L}$ e $\mathcal{M}$ são tais que

$$\dim\ \mathcal{N}(L) + \dim\ \mathcal{N}(M) < n.$$

Nesse caso, tais variedades dizem-se *enviesadas* se $\mathcal{L} \cap \mathcal{M} = \emptyset$ e $\mathcal{N}(L) \cap \mathcal{N}(M) = \{\vec{0}\}$, [4].

A inversa de *Moore-Penrose*[2] de uma matriz $M \in \mathbb{R}^{m\times n}$ é a única matriz $M^\dagger \in \mathbb{R}^{n\times m}$ satisfazendo

$$MM^\dagger M = M,\ M^\dagger MM^\dagger = M^\dagger,\ (M^\dagger M)^T = M^\dagger M \text{ e } (MM^\dagger)^T = MM^\dagger.$$

Em particular, se M é uma matriz com característica completa por linhas, então

$$M^\dagger = M^T \left(MM^T\right)^{-1}.$$

No que se segue, assumiremos que todas as matrizes envolvidas têm característica completa por linhas.

No final desta nota, exibiremos o melhor par aproximante $\left(\vec{x^*}, \vec{y^*}\right) \in \mathcal{L} \times \mathcal{M}$ de duas variedades lineares enviesadas $\mathcal{L}$ e $\mathcal{M}$, querendo isto dizer que

$$\left\|\vec{x^*} - \vec{y^*}\right\| = d\left(\mathcal{L}, \mathcal{M}\right),$$

onde $d\left(\mathcal{L}, \mathcal{M}\right)$ designa a distância euclidiana entre as variedades $\mathcal{L}$ e $\mathcal{M}$.

3 Projeção de um Ponto sobre uma Variedade Linear

Considere uma variedade linear

$$\mathcal{L} = \vec{l_0} + \mathcal{N}(L) = \{\vec{x} \in \mathbb{R}^n : L\vec{x} = \vec{c}\},$$

onde $\vec{l_0}$ é um ponto fixo de $\mathbb{R}^n$. Procura-se determinar $\mathbb{P}_{\mathcal{L}}\left(\vec{m}\right)$, a projeção (ortogonal) de um dado ponto externo $\vec{m} \in \mathbb{R}^n$ sobre $\mathcal{L}$, isto é, pretende-se determinar o ponto $\vec{l} \in \mathcal{L}$ que minimize a distância $d\left(\vec{l}, \vec{m}\right)$.

Para atingir o referido objetivo, propõe-se um método original que recorre duas vezes a um resultado que fornece o vetor de norma mínima de uma variedade linear com codimensão finita. Tal resultado pode ser consultado, *e.g.*, em [3, Theorem 9.26, p. 215], [8, Théorème 2.2.5, p. 45] e [9, Theorem 2, p. 51], assumindo a seguinte forma.

Teorema 3.1. *Considere uma variedade linear $\mathcal{L} = \vec{l_0} + \mathcal{N}(L)$ de $\mathbb{R}^n$. Uma condição necessária e suficiente para que $\vec{l} \in \mathcal{L}$ seja a melhor aproximação de $\vec{m} \notin \mathcal{L}$ em $\mathcal{L}$ é que $\vec{l} - \vec{m}$ seja ortogonal a $\mathcal{N}(L)$.*

Considerando esta caracterização do ponto que constitui a melhor aproximação de uma variedade linear, temos então o principal resultado desta secção.

[2] Na referência [1] é detalhada a teoria das inversas generalizadas de matrizes.

Teorema 3.2. *Sejam* $\mathcal{L} = \vec{l}_0 + \mathcal{N}(L)$ *uma variedade linear e* $\vec{m} \notin \mathcal{L}$ *um dado ponto fixo. A projeção* $\mathbb{P}_{\mathcal{L}}(\vec{m})$ *é dada por*

$$\mathbb{P}_{\mathcal{L}}(\vec{m}) = \vec{x}_0 + \vec{m} - \vec{x}'_0, \tag{1}$$

onde $\vec{x}_0$ *e* $\vec{x}'_0$ *são, respetivamente, os vetores de norma mínima das variedades lineares paralelas* $\mathcal{L}$ *e* $\mathcal{L}' = \vec{m} + \mathcal{N}(L)$.

Demonstração. Tendo em conta o Teorema 3.1, temos de provar que $\vec{m} - \mathbb{P}_{\mathcal{L}}(\vec{m})$ é ortogonal ao subespaço $\mathcal{N}(L)$.

De facto, tem-se

$$\mathbb{P}_{\mathcal{L}}(\vec{m}) - \vec{m} = \vec{x}_0 - \vec{x}'_0.$$

Mas $\vec{x}_0 - \vec{x}'_0$ é ortogonal a $\mathcal{N}(L)$ em virtude do facto de $\vec{x}_0$ e $\vec{x}'_0$ serem ambos ortogonais a $\mathcal{N}(L)$ (ver [9, Theorem 1, p. 64]). □

Teorema 3.3. *Considere uma variedade linear*

$$\mathcal{L} = \{\vec{x} \in \mathbb{R}^n : \ L\vec{x} = \vec{c}\}.$$

1. *O vetor de norma mínima de* $\mathcal{L}$ *é dado por* $\vec{x}_0 = L^\dagger \vec{c}$;
2. $\mathbb{P}_{\mathcal{N}(L)}(\vec{y}) = \left(I - L^\dagger L\right)\vec{y}$;
3. $\mathbb{P}_{\mathcal{L}}(\vec{y}) = \left(I - L^\dagger L\right)\vec{y} + L^\dagger \vec{c}$.

Demonstração. Relativamente a 1. e 2., consulte [10, pp. 423 e 434–435].

No que diz respeito a 3., seja $\vec{x}_0$ o vetor de norma mínima de $\mathcal{L}$. Por definição de projeção ortogonal, temos

$$\mathbb{P}_{\mathcal{L}}(\vec{y}) = \vec{x}_0 + \mathbb{P}_{\mathcal{N}(L)}(\vec{y} - \vec{x}_0),$$

de onde resulta

$$\mathbb{P}_{\mathcal{L}}(\vec{y}) = \left(I - \mathbb{P}_{\mathcal{N}(L)}\right)\vec{x}_0 + \mathbb{P}_{\mathcal{N}(L)}(\vec{y}). \tag{2}$$

Por 1. e 2. e pela definição de $L^\dagger$ tem-se:

$$\left(I - \mathbb{P}_{\mathcal{N}(L)}\right)\vec{x}_0 = \left[I - \left(I - L^\dagger L\right)\right]\vec{x}_0 = L^\dagger L \vec{x}_0 = L^\dagger L L^\dagger \vec{c} = L^\dagger \vec{c}.$$

A expressão em 3. obtém-se após substituição em (2) e aplicação de 2. □

Tendo em conta o resultado anterior, podemos agora exprimir o Teorema 3.2 à custa da inversa de Moore-Penrose aplicada à representação dos vetores de norma mínima das variedades lineares paralelas ali consideradas.

Corolário 3.4. *Nas condições do Teorema 3.2, sendo* $\mathcal{L} = \{\vec{x} \in \mathbb{R}^n : L\vec{x} = \vec{c}\}$ *a representação matricial da variedade* $\mathcal{L}$*, tem-se:*

$$\mathbb{P}_{\mathcal{L}}(\vec{m}) = \left(I - L^{\dagger}L\right)\vec{m} + L^{\dagger}\vec{c}. \tag{3}$$

Demonstração. De modo a exprimir os vetores de norma mínima em (1), vamos recorrer à notação matricial das variedades $\mathcal{L}$ e $\mathcal{L}^{'}$. Assim, dado que tais variedades são paralelas, e tendo em conta a representação de $\mathcal{L}$, podemos assumir que

$$\mathcal{L}^{'} = \{\vec{x} \in \mathbb{R}^n : L\vec{x} = \vec{d}\},$$

para determinados $\vec{d} \in \mathbb{R}^p$. Uma vez que $\mathcal{L}^{'}$ passa por $\vec{m}$, tem-se $L\vec{m} = \vec{d}$. Deste modo, por aplicação de 1. do Teorema 3.3 tem-se:

$$\vec{x}_0 = L^{\dagger}\vec{c} \quad \text{e} \quad \vec{x}_0^{'} = L^{\dagger}\vec{d} = L^{\dagger}L\vec{m}.$$

A expressão (3) obtém-se após substituição de $\vec{x}_0$ e $\vec{x}_0^{'}$ em (1), seguida de simplificação. □

4 Melhor Par Aproximante

Nesta última secção, procuramos obter o mais curto segmento de reta

$$\left[\vec{x^*}\ \vec{y^*}\right]$$

ligando duas variedades lineares enviesadas $\mathcal{L}$ e $\mathcal{M}$.

Para tal, aplicamos duas vezes o Teorema 3.2, levando em conta que, no presente caso, os pontos externos a considerar são pontos genéricos das variedades lineares $\mathcal{L}$ e $\mathcal{M}$, respetivamente.

Teorema 4.1. *Considere duas variedades lineares enviesadas,* $\mathcal{L} = \vec{l}_0 + \mathcal{N}(L)$ *e* $\mathcal{M} = \vec{m}_0 + \mathcal{N}(M)$*. Então, o melhor par aproximante* $\left(\vec{x^*},\ \vec{y^*}\right) \in \mathcal{L} \times \mathcal{M}$ *das variedades lineares* $\mathcal{L}$ *e* $\mathcal{M}$ *é a solução do sistema de equações lineares*

$$\left\{\begin{array}{l} \mathbb{P}_{\mathcal{L}}(\vec{g}_{\mathcal{M}}) = \vec{g}_{\mathcal{L}} \\ \mathbb{P}_{\mathcal{M}}(\vec{g}_{\mathcal{L}}) = \vec{g}_{\mathcal{M}} \end{array}\right. ,$$

onde $\vec{g}_{\mathcal{L}}$ *e* $\vec{g}_{\mathcal{M}}$ *são pontos genéricos pertencentes às variedades lineares* $\mathcal{L}$ *e* $\mathcal{M}$*, respetivamente.*

Demonstração. Comece-se por aplicar duas vezes o Teorema 3.2, uma vez que o ponto genérico $\vec{g}_{\mathcal{L}}$ é externo a $\mathcal{M}$ e que o ponto genérico $\vec{g}_{\mathcal{M}}$ é externo a $\mathcal{L}$. Depois, basta ter em conta que o vetor $\overrightarrow{\vec{x^*}\vec{y^*}}$ — cujas extremidades são $\vec{x^*} \in \mathcal{L}$ e $\vec{y^*} \in \mathcal{M}$ — é simultaneamente ortogonal aos subespaços $\mathcal{N}(L)$ e $\mathcal{N}(M)$. $\square$

Para terminar, vamos exibir o melhor par aproximante obtido no teorema anterior em que os pontos se exprimem à custa da inversa de Moore-Penrose.

Corolário 4.2. *Nas condições do Teorema 4.1, admita que as variedades lineares $\mathcal{L}$ e $\mathcal{M}$ são dadas na forma matricial do seguinte modo:*

$$\mathcal{L} = \{\vec{x} \in \mathbb{R}^n : L\vec{x} = \vec{c}\} \text{ e } \mathcal{M} = \{\vec{x} \in \mathbb{R}^n : M\vec{x} = \vec{d}\}.$$

Então o melhor par aproximante $\left(\vec{x^*},\ \vec{y^*}\right) \in \mathcal{L} \times \mathcal{M}$ é a solução do sistema

$$\begin{cases} \vec{x} = \left(I - L^\dagger L\right)\vec{y} + L^\dagger \vec{c} \\ \vec{y} = \left(I - M^\dagger M\right)\vec{x} + M^\dagger \vec{d}. \end{cases}$$

Demonstração. Sendo $\vec{x} \in \mathcal{L}$ e $\vec{y} \in \mathcal{M}$ tais que

$$\begin{cases} \vec{x} = \mathbb{P}_{\mathcal{L}}(\vec{y}) \\ \vec{y} = \mathbb{P}_{\mathcal{M}}(\vec{x}), \end{cases}$$

resulta de imediato por aplicação do Corolário 3.4 que

$$\begin{cases} \vec{x} = \left(I - L^\dagger L\right)\vec{y} + L^\dagger \vec{c} \\ \vec{y} = \left(I - M^\dagger M\right)\vec{x} + M^\dagger \vec{d}. \end{cases}$$

$\square$

Agradecimentos

Os autores agradecem ao revisor, cujas importantes observações os levaram a refazer o artigo, tanto na forma como no conteúdo.

Referências

[1] A. Ben-Israel e T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Springer-Verlag, New York, 2003.

[2] A. Dax, "The Distance between Two Convex Sets", *Linear Algebra and its Applications*, Vol. 416, No. 1 (2006), pp. 184–213. DOI: `https://doi.org/10.1016/j.laa.2006.03.022`

[3] F. Deutsch, Best *Approximation in Inner Product Spaces*, CMS books in mathematics, 7, Springer-Verlag, New York, 2001.

[4] A. M. DuPré e S. Kass, "Distance and Parallelism between Flats in $\mathbb{R}^n$", *Linear Algebra and its Applications*, Vol. 171 (1992), pp. 99-107. DOI: `https://doi.org/10.1016/0024-3795(92)90252-6`

[5] M. A. Facas Vicente, Armando Gonçalves e José Vitória, "Euclidean distance between two linear varieties", *Applied Mathematical Sciences*, Vol. 8, No. 21 (2014), pp. 1039-1043. DOI: `http://dx.doi.org/10.12988/ams.2014.311656`

[6] Armando Gonçalves, M. A. Facas Vicente e José Vitória, "Optimal pair of two linear varieties", *Applied Mathematical Sciences*, Vol. 9, No. 12 (2015), pp. 593-596. DOI: `http://dx.doi.org/10.12988/ams.2015.410869`

[7] J. Gross e G. Trenkler, "On the Least Squares Distances between Affine Subspaces", *Linear Algebra and its Applications*, Vols. 237/238 (1996), pp. 269-276. DOI: `http://dx.doi.org/10.1016/0024-3795(95)00648-6`

[8] P.-J. Laurent, *Approximation et Optimisation*, Hermann, Paris, 1972.

[9] D. G. Luenberger, *Optimization by Vector Space Methods*, J. Wiley, New York, 1969.

[10] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[11] G. Wang, Yimin Wei e Sanzheng Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing, New York, 2004.