Uma Aplicação das Funções de Matrizes na Óptica Linear

João R. Cardoso
Instituto Superior de Engenharia de Coimbra
Rua Pedro Nunes,
3030-199 Coimbra, Portugal
e-mail: `jocar@isec.pt`

William F. Harris
Department of Optometry
University of Johannesburg
APK Campus, P. O. Box 524,
Auckland Park, 2006, África do Sul
e-mail: `wharris@uj.ac.za`

Resumo: Em óptica linear, um sistema óptico fica completamente caracterizado por uma matriz 5×5 com uma estrutura especial, denominada de matriz de transferência dos raios. Alguns dos exemplos mais correntes de sistemas ópticos incluem o olho humano, as lentes e os telescópios. Uma questão importante em óptica linear é o cálculo da média de um conjunto de sistemas ópticos. No caso do olho humano, isto corresponde a determinar o olho padrão de uma população. Matematicamente, o que está em causa é o cálculo da média de um conjunto de matrizes, em que determinadas condições têm de ser satisfeitas. Em particular, a matriz correspondente à média obtida terá de ser uma matriz de transferência de raios. Neste artigo iremos recordar algum do material incluído em [*Ophthal. Physiol. Optics*, 26 (2006), pp. 380–383], onde é proposta uma solução para este problema baseada em funções de matrizes bem conhecidas, como é o caso da exponencial e do logaritmo de matrizes. Para facilitar a compreensão do problema, será apresentado um exemplo prático e serão recordados diversos conceitos e resultados relacionados.

Abstract In linear optics an optical system (e.g., an eye, a lens, a telescope) is completely characterized by a 5×5 matrix with a special structure, called the ray transference. An important issue in the quantitative analysis of optical systems is the question of how to calculate an average of a set of eyes or other optical systems. Mathematically, this corresponds to averaging a set of matrices (ray transferences), providing that some requirements are fulfilled. In particular, the average of ray transferences must be a ray transference. In this paper we revisit some of the material included in [*Ophthal.*

Physiol. Optics, 26 (2006), pp. 380–383], where a solution for this problem based on well-known matrix functions such as the matrix exponential and the matrix logarithm is proposed . To ease the understanding of the problem, a practical example is presented as well as some related concepts and results are recalled.

palavras-chave: Óptica linear, matriz de transferência dos raios, matriz simplética, exponencial de matrizes, logaritmo de matrizes.

keywords: Linear optics, ray transference matrix, symplectic matrix, matrix exponential, matrix logarithm.

1 Introdução

As primeiras referências a funções de matrizes são datadas dos finais do século XIX e estão incluídas em trabalhos de matemáticos famosos como Arthur Cayley e James Sylvester. Para o caso particular da exponencial de matrizes, também são conhecidas referências em trabalhos de Edmond Laguerre e Giuseppe Peano, datados da mesma época. No entanto, é nas décadas de 60 e 70 do século XX que se assiste ao aparecimento de um grande número de trabalhos científicos sobre este tópico, a maioria deles propondo novos métodos para o cálculo destas funções. A este interesse crescente não é alheia a existência de uma grande diversidade de aplicações em várias áreas do conhecimento, como, por exemplo, em Teoria do Controlo, Física, Sociologia e Engenharia. Uma das aplicações recentes que tem despertado algum interesse é precisamente aquela que é objecto de estudo neste artigo (ver, em particular, as referências a esta aplicação em [5] e [12, Sec. 2.14.5]). Trata-se de uma aplicação a um problema de Óptica Linear que, em linhas gerais, consiste em determinar a matriz média $\overline{T}$ de um conjunto de N matrizes de transferências de raios $T_1, T_2, \ldots, T_N$, representando cada uma delas um certo sistema óptico. É requerido que a matriz média satisfaça um conjunto de propriedades, não verificadas em geral pelas médias mais usuais, nomeadamente, a média aritmética e a média geométrica. Como forma de contornar esta dificuldade, Harris [9] conjectura que a chamada média `exp-log`,

$$\overline{T} = \exp\left(\frac{1}{N}\sum_{j=1}^{N} \log T_j\right),$$

verifica as propriedades desejadas. Esta conjectura foi estudada em [11], onde se mostrou ser verdadeira, desde que sejam impostas algumas restrições aos valores próprios das matrizes de transferências de raios.

Para melhor entendermos a problemática do cálculo de um "sistema óptico médio", iremos começar por rever alguns conceitos importantes para a compreensão deste tipo de sistemas. Assim, na Secção 2, faremos uma breve introdução à Óptica Linear, também designada por Óptica Paraxial, onde serão revistos alguns conceitos básicos e também as equações que servem de modelo a um sistema óptico. Nas secções 3 e 4 são consideradas duas matrizes importantes no estudo de sistemas ópticos em óptica linear: a matriz da potência dióptrica e a matriz de transferência dos raios. Um exemplo prático de matrizes de potência dióptrica associadas a um par de lentes de óculos convencionais é apresentado. As propriedades requeridas para a média de um conjunto de matrizes de transferência de raios são apresentadas na Secção 5, em que será definido o conceito de média significativa. As definições de exponencial e logaritmo de matrizes serão recordadas na Secção 6, bem como algumas propriedades e referências bibliográficas para um estudo mais aprofundado do assunto. Na Secção 7 será apresentado um teorema contendo o resultado central deste artigo, o qual mostra que, mediante certas restrições no espectro das matrizes de transferências de raios envolvidas, `exp-log` é uma média significativa, podendo portanto ser usada no cálculo do sistema óptico médio. Um exemplo ilustrativo da utilização da média `exp-log` no cálculo do olho padrão será apresentado.

2 Breves Noções de Óptica Linear

De uma forma abreviada, pode-se dizer que a óptica é um ramo da física que se ocupa essencialmente do estudo da luz. Tal como na física em geral, também na óptica a matemática desempenha um papel fundamental na compreensão dos fenómenos. No caso da óptica geométrica é a geometria a ferramenta mais utilizada, enquanto que na óptica linear, também designada por alguns autores como óptica paraxial ou óptica de primeira ordem, se utilizam com bastante frequência os métodos matriciais e outras ferramentas da álgebra linear. Outros ramos da óptica não considerados neste trabalho incluem a óptica ondulatória, óptica electromagnética ou óptica quântica.

A óptica linear é uma aproximação da óptica geométrica, na medida em que se assume que os ângulos envolvidos têm uma amplitude pequena. É natural considerar como válidas as aproximações

$$\sin\theta \approx \theta, \quad \tan\theta \approx \theta, \quad \cos\theta \approx 1,$$

onde θ representa a amplitude de um certo ângulo. Para exemplificar, consideremos uma das leis mais importantes em óptica geométrica que é a lei

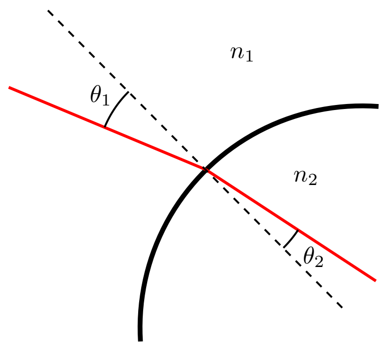


Figura 1: Ilustração de lei de Snell

de Snell. Esta lei estabelece que se um raio de luz passa de uma região com índice de refracção n_1 para uma região com índice de refracção n_2, então

$$n_1 \sin\theta_1 = n_2 \sin\theta_2,$$

onde θ_1 e θ_2 são os ângulos que o raio de luz faz com a normal à superfície que separa as duas regiões (Figura 1). Em óptica linear, atendendo a que os ângulos têm uma amplitude pequena, esta lei é traduzida pela equação bem mais simples,

$$n_1\,\theta_1 = n_2\,\theta_2$$

em que as funções trigonométricas admitem uma representação linear em função dos ângulos. Recorde-se que esta aproximação da lei de Snell já tinha sido usada muito antes por Ptolomeu.

Uma versão simplificada da óptica linear é a chamada Óptica Gaussiana[1] em que se assume que as superfícies em consideração são planas ou parte de uma esfera. Este ramo da óptica permite a utilização de ferramentas matemáticas simples no estudo de alguns sistemas ópticos importantes, como por exemplo as lentes delgadas, espelhos e também o olho humano, desde que o astigmatismo seja tão pequeno que possa ser desprezado. Do ponto de vista geométrico, em óptica Gaussiana, um sistema óptico pode ser globalmente representado num plano, habitualmente designado por plano zOy, em que o eixo horizontal Oz é designado por *eixo óptico*. Um raio de luz é representado localmente por um vector, como se exemplifica a seguir.

Suponhamos que um raio luminoso se desloca num meio com índice de refracção n. Quando este passa num determinado ponto, fica caracterizado

[1] Esta designação é devida às importantes contribuições do famoso matemático e físico alemão Carl Friedrich Gauss (1777–1855) no estudo dos sistemas ópticos

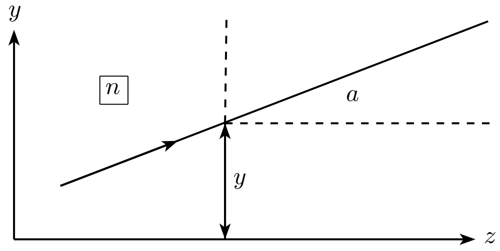


Figura 2: Coordenadas de um raio de luz num ponto, ao deslocar-se num meio com índice de refracção n. y é a distância do ponto ao eixo óptico e a o ângulo formado pela trajectória do raio com a horizontal.

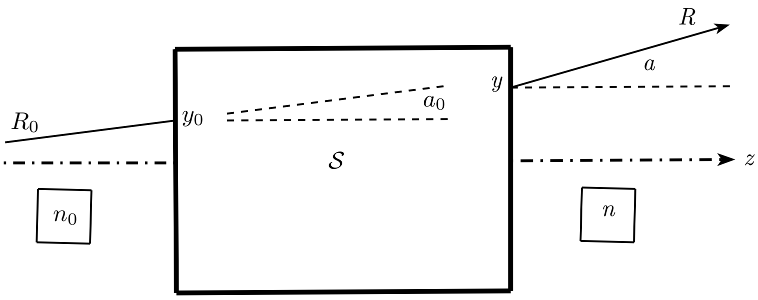


Figura 3: O raio R_0R atravessa o sistema óptico $\mathcal{S}$ em óptica Gaussiana. z é o eixo óptico. O raio entra no sistema na posição y_0 e ângulo de inclinação a_0 e sai do sistema na posição y e ângulo a. Os índices de refracção dos meios anterior e posterior ao sistema são respectivamente n_0 e n.

pela distância y do ponto ao eixo óptico e pelo ângulo a formado pela sua trajectória (rectilínea) com o eixo óptico Oz (Figura 2). Por uma questão de simplificação de algumas das equações matemáticas envolvidas, a coordenada de orientação a é substituída habitualmente pela chamada coordenada de orientação óptica $\alpha = n\,a$. Assim, ao passar num ponto, um raio fica caracterizado pelo vector 2×1

$$\begin{bmatrix} y \\ \alpha \end{bmatrix}.$$

Estas coordenadas podem ser alteradas quando o raio atravessa um sistema óptico. Do ponto de vista geométrico, um raio pode ser interpretado como uma sequência de vectores. É frequente designar o vector representativo do raio que entra num sistema óptico por segmento ou raio incidente e o correspondente vector de saída por segmento ou raio emergente. A Figura 3 mostra os segmentos incidente R_0 e emergente R de um raio que atravessa um sistema óptico $\mathcal{S}$ em óptica Gaussiana. Suponhamos que o segmento incidente entra no sistema na posição y_0 com inclinação a_0. As coordenadas do vector representativo do raio na entrada são

$$\rho_0 = \begin{bmatrix} y_0 \\ \alpha_0 \end{bmatrix},$$

em que $\alpha_0 = n_0 \, a_0$, sendo n_0 o índice de refracção do meio anterior ao sistema. De forma análoga, vamos supor que o raio emergente é definido pelo vector

$$\rho = \begin{bmatrix} y \\ \alpha \end{bmatrix},$$

em que $\alpha = na$. Assume-se que os ângulos de incidência e emergência são pequenos. A equação básica deste sistema em óptica Gaussiana é

$$S \, \rho_0 = \rho,$$

onde S é uma certa matriz 2×2 simplética, denominada por *matriz da transferência dos raios* (ver Secção 4). As entradas de S dependem do tipo de sistema em consideração. Tal como foi referido acima, a óptica Gaussiana pode ser usada para estudar um sistema, desde que o seu astigmatismo possa ser considerado desprezível. Se tal não acontecer, é necessário passar para o contexto da óptica linear. Neste caso, os raios incidente ρ_0 e emergente ρ são caracterizados por vectores 4×1 e a matriz simplética que caracteriza o sistema óptico terá de ser 4×4.

Se o leitor estiver interessado em mais detalhes sobre óptica Gaussiana, óptica geométrica e óptica linear, bem como sobre a relação entre elas, a referência [3, cap. IX] proporciona uma abordagem interessante. Em particular, nesta referência, e também em [1], poderá ser encontrada a dedução da expressão das entradas das matrizes simpléticas que caracterizam os sistemas ópticos mais conhecidos.

3 Matriz da Potência Dióptrica

Numa prescrição de lentes para óculos, o oftalmologista (ou optometrista) precisa de indicar os valores de três componentes fundamentais da lente: a potência esférica (F_S), a potência cilíndrica (F_C) e o ângulo de inclinação do eixo cilíndrico (A). É frequente estas três componentes aparecerem escritas de acordo com a notação

$$F_S/\ F_C\ \times\ A. \tag{1}$$

A unidade referente a F_S e a F_C é a dioptria (D) e A é medido em graus; $1\,D$ equivale a $1\,m^{-1}$, em que m designa o metro. A leitura destes três números é feita de um modo particular. Por exemplo,

$$+4/\ -2 \times 30$$

lê-se "mais quatro, menos dois eixo 30". À prescrição (1) está associada a matriz simétrica, 2×2,

$$F = \begin{bmatrix} F_S + F_C\ \sin^2 A & -F_C \sin A\ \cos A \\ -F_C \sin A\ \cos A & F_S + F_C\ \cos^2 A \end{bmatrix},$$

à qual se dá o nome de *matriz da potência dióptrica* [15, 14]. Não é difícil mostrar que os valores próprios de F são

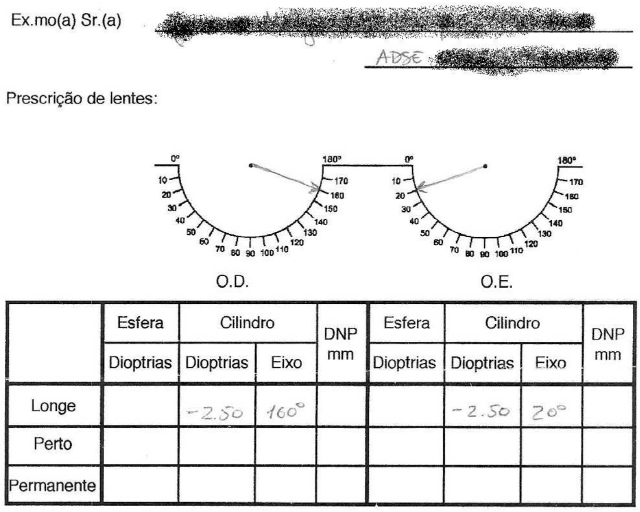
$$\lambda_1 = F_S, \quad \lambda_2 = F_S + F_C,$$

aos quais estão associados, respectivamente, os vectores próprios

$$\mathbf{u}_1 = \begin{bmatrix} \cos A \\ \sin A \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} \sin A \\ -\cos A \end{bmatrix}.$$

Verificamos assim que os valores e vectores próprios da matriz F estão directamente relacionados com a prescrição. Note-se que o ângulo de inclinação do vector próprio associado à potência esférica coincide precisamente com o ângulo indicado em (1). A representação das características de uma lente através de uma matriz simétrica tem inúmeras vantagens, tal como é reportado em [8]. Uma delas diz respeito ao facto de a representação matricial permitir um tratamento matemático mais fácil dos assuntos. Além disso, para uma mesma lente podem existir várias prescrições da forma (1), o que pode originar alguns problemas. No entanto, cada lente tem apenas uma matriz da potência dióptrica.

Na Figura 4 podemos observar uma prescrição de lentes efectuada recentemente. Hoje em dia, é frequente substituir-se a notação (1) pelo preenchimento de uma tabela onde são colocados os valores das três componentes da lente. Quando não é indicado qualquer valor para a potência esférica, assume-se que é nula. Usando quatro casas decimais correctas, as matrizes



Ex.mo(a) Sr.(a) [redacted]

ADSE [redacted]

Prescrição de lentes:

O.D. O.E.

	Esfera	Cilindro		DNP mm	Esfera	Cilindro		DNP mm
	Dioptrias	Dioptrias	Eixo		Dioptrias	Dioptrias	Eixo	
Longe		-2.50	160°			-2.50	20°	
Perto								
Permanente								

Figura 4: Exemplo de uma prescrição de lentes

da potência dióptrica das lentes do olho direito F_{OD} (OD de "oculus dexter") e do olho esquerdo F_{OS} (OS de "oculus sinister") são, respectivamente,

$$F_{OD} = \begin{bmatrix} -0.2924 & -0.8035 \\ -0.8035 & -2.2076 \end{bmatrix}, \quad F_{OS} = \begin{bmatrix} -0.2924 & 0.8035 \\ 0.8035 & -2.2076 \end{bmatrix}.$$

Note-se que as duas matrizes apenas diferem nos sinais das entradas não diagonais. Tal está de acordo com a simetria bilateral do corpo humano relativamente ao plano médio sagital.

4 Matriz de Transferência dos Raios

Para o caso de sistemas ópticos mais complexos que as lentes de óculos (como, por exemplo, o olho humano), são necessários diversos parâmetros para os caracterizar. Em óptica linear, qualquer sistema óptico fica completamente caracterizado por uma matriz 5×5, denominada por matriz de transferência dos raios [7]. Antes de definir esta matriz, convém recordar o que é uma matriz simplética.

Denotemos por $\mathbb{R}^{m\times m}$ o conjunto das matrizes quadradas $m \times m$ com entradas em $\mathbb{R}$, I a matriz identidade e 0 a matriz nula. Considere-se a matriz em blocos

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{2}$$

Diz-se que uma matriz $S \in \mathbb{R}^{2m\times 2m}$ é *simplética* se $S^T JS = J$, onde S^T designa a transposta de S. Se particionarmos uma matriz simplética S em blocos

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

em que A, B, C, $D \in \mathbb{R}^{m\times m}$, é fácil mostrar que estes verificam as relações

$$\begin{aligned} A^T C &= C^T A \\ B^T D &= D^T B \\ A^T D - C^T B &= I. \end{aligned}$$

Chama-se *matriz de transferência dos raios* a uma matriz 5×5 particionada em blocos na forma:

$$T = \begin{bmatrix} A & B & \delta_1 \\ C & D & \delta_2 \\ 0 & 0 & 1 \end{bmatrix},$$

onde a submatriz 4×4 que ocupa a parte superior esquerda de T,

$$S = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

é simplética. Os blocos que formam a matriz T, A, B, C, $D \in \mathbb{R}^{2\times 2}$ e δ_1, $\delta_2 \in \mathbb{R}^{2\times 1}$, contêm toda a informação referente às seis propriedades fundamentais que caracterizam qualquer sistema óptico. Convém ainda notar que o bloco C está directamente relacionado com a matriz da potência dióptrica do sistema, F, na medida em que $F = -C$.

Suponha-se que um dado sistema óptico $\mathcal{S}$ corresponde à combinação de N sucessivos sistemas ópticos $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_N$, com matrizes de transferência dos raios, respectivamente, $T_1, T_2, \ldots, T_N$. Eis uma das vantagens de representar matricialmente sistemas ópticos: a matriz de transferência de raios de $\mathcal{S}$, T, é o produto, por ordem inversa, das matrizes de transferência que compõem os sistema, isto é,

$$T = T_N T_{N-1} \ldots T_1.$$

5 Média de Matrizes de Transferência de Raios

Um dos problemas típicos em óptica linear é a caracterização do olho padrão de uma determinada população. Uma vez que um olho pode ser caracterizado por uma matriz de transferência dos raios, matematicamente, tal corresponde a calcular a média das matrizes de transferência correspondentes. Existem várias formas de calcular médias de matrizes, mas temos de ter em conta que a matriz obtida para caracterizar o olho padrão tem de ser, em particular, uma matriz de transferência dos raios. Há ainda outras condições que são exigidas, como se mostra na definição seguinte.

Definição. Dadas as matrizes de transferência de raios $T_1, T_2, \ldots, T_N$, diz-se que uma aplicação

$$(T_1, T_2, \ldots, T_N) \longmapsto T = \Phi(T_1, T_2, \ldots, T_N)$$

é uma *média significativa* se forem satisfeitas as três condições seguintes:

(i) T é uma matriz de transferência dos raios;

(ii) Φ é independente da ordem considerada para as matrizes T_j, $j = 1, \ldots, N$;

(iii) Φ é invariante relativamente à mudança de unidades.

Aqui é necessário explicar o significado da condição (iii). As componentes de uma matriz de transferência de raios têm significados físicos diversos aos quais correspondem unidades distintas. É necessário que o processo de cálculo da média não dependa do sistema de unidades que se está a usar, pelo que se exige a invariância relativamente às unidades consideradas. Matematicamente, esta invariância corresponde à preservação de uma semelhança de matrizes, como se explica de seguida.

Para cada $j = 1, \ldots, N$, consideremos a parte simplética de cada matriz T_j,

$$S_j = \begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix}$$

e suponhamos que a média destas matrizes é

$$\overline{S} = \Phi(S_1, S_2, \ldots, S_N).$$

Dado um número real positivo β, defina-se a matriz 4×4,

$$D = \begin{bmatrix} \frac{1}{\sqrt{\beta}}\, I & 0 \\ 0 & \sqrt{\beta}\, I \end{bmatrix}.$$

Então $\Phi(S_1, S_2, \ldots, S_N)$ será invariante relativamente à mudança das unidades se

$$\Phi(DS_1D^{-1}, DS_2D^{-1}, \ldots, DS_ND^{-1}) = D\,\Phi(S_1, S_2, \ldots, S_N)\,D^{-1}.$$

Convém notar que nem a média aritmética

$$\overline{T} = \frac{T_1 + T_2 + \ldots + T_N}{N}$$

nem a média geométrica

$$\widetilde{T} = (T_1 T_2 \cdots T_N)^{1/N}$$

são médias significativas de matrizes de transferência dos raios. A primeira não verifica a condição (i) da definição — a soma de matrizes simpléticas não é em geral simplética — e a segunda média não verifica a condição (ii), uma vez que o produto de matrizes não é em geral comutativo.

Antes do aparecimento dos computadores, o cálculo da média geométrica de N números reais positivos $\alpha_1, \alpha_2, \ldots, \alpha_N$ era, por vezes, feito com base nas famosas tabelas de logaritmos. Para tal, em vez da fórmula usual $\overline{\alpha} = (\alpha_1\, \alpha_2 \ldots \alpha_N)^{1/N}$ usava-se a expressão equivalente

$$\overline{\alpha} = \exp\left(\frac{1}{N}\sum_{j=1}^{N} \log \alpha_j\right).$$

Foi nesta ideia que Harris [9] se baseou para conjecturar, com base em diversos exemplos práticos, que a chamada média `exp-log`

$$\overline{T} = \exp\left(\frac{1}{N}\sum_{j=1}^{N} \log T_j\right),$$

onde exp(.) e log(.) denotam, respectivamente, a exponencial e o logaritmo principal de matrizes, verificava as condições que definem uma média significativa. Para mostrar a veracidade desta conjectura, é necessário usar ferramentas da teoria das funções de matrizes.

6 Exponencial e Logaritmo de Matrizes

Nesta secção faremos apenas uma breve revisão sobre alguns aspectos relevantes das funções exponencial e logaritmo de matrizes. Para uma informação mais detalhada sobre este assunto e sobre a teoria geral das funções de matrizes remetemos o leitor para [12]; as referências [6, 13] também têm capítulos dedicados a este tema; para o caso particular do logaritmo ver [4].

A exponencial de matrizes é definida para uma matriz $A \in \mathbb{R}^{m\times m}$ através da seguinte série de potências com raio de convergência infinito:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2!} + \cdots .$$

Dada uma matriz $A \in \mathbb{R}^{n\times n}$, a qualquer matriz X que verifique a equação matricial

$$\exp(X) = A$$

dá-se o nome de logaritmo de A. Para uma matriz A ter pelo menos um logaritmo (que pode ser uma matriz complexa) é necessário e suficiente que A seja não singular. Em geral, uma matriz não singular admite uma infinidade de logaritmos. No entanto, se A não tiver valores próprios em $\mathbb{R}_0^-$ (o conjunto dos números reais não positivos), é garantida a existência de um único logaritmo de A com espectro contido na faixa do plano complexo definida por $\{z \in \mathbb{C} : \; -\pi < \operatorname{Im} z < \pi\}$; $\operatorname{Im} z$ designa a parte imaginária do número complexo z. Este único logaritmo é designado por *logaritmo principal* de A, que se representa por $\log A$ e é frequentemente designado, por abuso de linguagem, como sendo o logaritmo de A, omitindo-se o termo "principal". Algum software de cálculo (por exemplo, o MATHEMATICA, o MATLAB ou o MAPLE) inclui rotinas para calcular quer a exponencial quer o logaritmo de matrizes. No caso do MATLAB, os comandos usados são, respectivamente, `expm` e `logm`.

Existem outras definições alternativas para a exponencial e logaritmo de matrizes. Podem ser definidos, por exemplo, através da forma canónica de Jordan, do polinómio interpolador de Hermite ou da fórmula integral de Cauchy.

Uma matriz $H \in \mathbb{R}^{2m\times 2m}$ diz-se Hamiltoniana se $H^T J = -J\,H$, onde J é a matriz (2). As funções exponencial e logaritmo de matrizes relacionam as matrizes Hamiltonianas com as simpléticas. Com efeito, se H é Hamiltoniana então $\exp(H)$ é simplética; reciprocamente, se S é simplética com valores próprios não pertencentes a $\mathbb{R}_0^-$, então $\log S$ é simplética.

7 A Média exp-log

Teorema. [11] Dadas N matrizes de transferência de raios, $T_1, \ldots, T_N$, com valores próprios não pertencentes a $\mathbb{R}_0^-$, a média `exp-log`

$$\overline{T} = \exp\left(\frac{1}{N}\sum_{j=1}^{N}\log T_j\right),$$

é uma média significativa.

Demonstração. É imediato verificar que a média `exp-log` verifica a condição (ii) da definição de média significativa. O mesmo acontece com a condição (iii), se atendermos a que as funções de matrizes preservam as transformações de semelhança: $f(PBP^{-1}) = P\,f(B)\,P^{-1}$, para qualquer matriz não singular P [12, Teo. 1.13]. Resta-nos mostrar que a condição (i) é satisfeita.

Consideremos uma matriz de transferência dos raios T, com valores próprios não contidos em $\mathbb{R}_0^-$, representada na forma

$$T = \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix},$$

onde S é a correspondente matriz simplética 4×4 e $\delta \in \mathbb{R}^{4\times 1}$. Vamos mostrar que a matriz

$$L := \begin{bmatrix} \log S & (\log S)(\log S - I)^{-1}\delta \\ 0 & 0 \end{bmatrix}$$

verifica $L = \log T$. Com efeito, alguns cálculos permitem concluir que exponencial de L é

$$\begin{aligned}\exp(L) &= \sum_{k=0}^{\infty}\frac{L^k}{k!} = \begin{bmatrix} \exp(\log S) & (\exp(\log S) - I)(\exp(\log S) - I)^{-1}\delta \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix} = T,\end{aligned}$$

ficando provado que L é um logaritmo de T. Resta mostrar que L é de facto o logaritmo principal, isto é, que os seus valores próprios estão situados na faixa do plano complexo definida por $-\pi < \operatorname{Im} z < \pi$. Ora, isto é evidente porque o espectro de L é constituído exactamente por 0 e pelos valores próprios de $\log S$, que, por definição de logaritmo principal, estão situados naquela faixa.

Para cada $j = 1, \ldots, N$, escreva-se

$$T_j = \begin{bmatrix} S_j & \delta_j \\ 0 & 1 \end{bmatrix}.$$

Em virtude do que foi estabelecido acima, podemos dizer que

$$\frac{1}{N}\sum_{j=1}^{N} \log T_j = \begin{bmatrix} \frac{1}{N}\sum_{j=1}^{N} \log S_j & \frac{1}{N}\sum_{j=1}^{N} (\log S_j)(S_j - I)^{-1}\delta_j \\ 0 & 0 \end{bmatrix}. \tag{3}$$

Atendendo a que $\log S_j$ é uma matriz Hamiltoniana, a soma de matrizes Hamiltonianas ainda é Hamiltoniana e a exponencial de uma matriz Hamiltoniana é simplética, conclui-se que a exponencial da matriz (3), que é precisamente $\bar{T}$, é uma matriz de transferência dos raios. ∎

Como se pode observar no teorema anterior, para se usar a média `exp-log` é necessário garantir que nenhuma das matrizes $T_1, \ldots, T_N$ tem valores próprios em $\mathbb{R}_0^-$. Isto traz algumas desvantagens, na medida em que não será possível calcular a média `exp-log` de certos sistemas ópticos. Um exemplo, são os telescópios Keplerianos. Acontece, porém, que as matrizes de transferência de certos telescópios Keplerianos têm todos os valores próprios reais negativos e, neste caso, um raciocínio análogo ao usado na demonstração do teorema anterior permite concluir que a variante da média `exp-log` definida por

$$\widetilde{T} := -\exp\left(\frac{1}{N}\sum_{j=1}^{N} \log(-T_j)\right),$$

corresponde a uma média significativa (ver [10]). No caso do olho humano, tanto quanto sabemos, na literatura não é conhecido qualquer exemplo em que a matriz de transferência de raios tenha valores próprios reais negativos. Mesmo em casos relativos a olhos com características extremas. Em geral, os seus valores próprios são complexos e têm a parte real positiva. Note-se que as entradas de uma matriz de transferência de raios associada ao olho humano não podem ter valores arbitrários, uma vez que têm de ter um significado preciso do ponto de vista óptico. Em síntese, de acordo com a nossa experiência, podemos dizer que apenas alguns casos pontuais de sistemas ópticos têm matrizes de transferência com valores próprios reais negativos.

Em [10] são analisados os valores próprios das matrizes de transferência para o caso particular de sistemas ópticos Gaussianos.

Seguidamente, vamos ilustrar o resultado do teorema com um exemplo numérico, onde serão usados valores aproximados com 4 casas decimais correctas. Consideremos as seguintes três matrizes de transferência de raios

$$\begin{aligned} T_1 &= \begin{bmatrix} 0 & 0 & 16.6832 & 0 \\ 0 & 0 & 0 & 16.6832 \\ -0.0599 & 0 & 0.9044 & 0 \\ 0 & -0.0599 & 0 & 0.9044 \end{bmatrix}, \\ T_2 &= \begin{bmatrix} 0.0639 & -0.0024 & 16.7290 & -0.0281 \\ -0.0027 & -0.1045 & -0.0283 & 16.5626 \\ -0.0564 & -0.0002 & 0.8871 & -0.0029 \\ -0.0002 & -0.0659 & -0.0030 & 0.8733 \end{bmatrix}, \\ T_3 &= \begin{bmatrix} 0.0340 & 0.0936 & 17.5127 & 0.1154 \\ 0.0940 & -0.0970 & 0.1159 & 17.4162 \\ -0.0554 & 0.0050 & 0.8614 & 0.0097 \\ 0.0050 & -0.0622 & 0.0098 & 0.8559 \end{bmatrix}, \end{aligned}$$

cada uma delas correspondendo a um olho humano. Como a linha 5 e a coluna 5 são triviais (isto é, são formadas por zeros, com excepção da entrada (5,5) que é 1), foi considerada apenas a parte simplética de cada transferência. Em cada uma das matrizes, as entradas do bloco 2×2 da parte superior direita estão em milímetros mm e as do bloco 2×2 da parte inferior esquerda em mm^{-1}. Por definição, às entradas dos dois blocos 2×2 da diagonal não correspondem quaisquer unidades de medidas. A transferência T_1 diz respeito ao olho completo "Le Grand", o qual foi obtido a partir da estrutura do olho providenciada em [2, p. 251]. Os quatro blocos 2×2 de T_1 são matrizes escalares, em virtude de o astigmatismo deste olho poder ser considerado nulo. Em particular, o bloco superior esquerdo é a matriz nula, devido ao facto de ser emétrope. Este olho não necessita de qualquer lente de compensação para ver correctamente objectos distantes. As transferências T_2 e T_3 correspondem a olhos amétropes, uma vez que os blocos situados na parte superior esquerda são não nulos. Ambos os olhos requerem lentes para ver de forma clara objectos distantes e têm astigmatismo. Devido à existência de astigmatismo, os quatro blocos de cada transferência não são matrizes escalares. Calculando a média `exp-log` no MATLAB, obtém-se

$$\overline{T} = \begin{bmatrix} 0.0317 & 0.0300 & 16.9731 & 0.0164 \\ 0.0300 & -0.0681 & 0.0168 & 16.8935 \\ -0.0573 & 0.0016 & 0.8837 & 0.0012 \\ 0.0016 & -0.0627 & 0.0012 & 0.8780 \end{bmatrix}.$$

A partir das entradas de $\overline{T}$ podem-se inferir diversas características específicas do olho médio, como por exemplo, o tamanho da imagem de um objecto na retina e a compensação do erro refractivo.

Agradecimentos: Os autores agradecem aos revisores anónimos e a Fátima Leite pelas sugestões e correcções propostas.

Referências

[1] J. Almeida, *Óptica Geométrica*, Departamento de Física, Universidade do Minho, 2008. `http://www.arauto.uminho.pt/pessoas/bda/opgeom.html`

[2] D. Atchison, G. Smith, *Optics of the Human Eye*, Butterworth-Heinemann, Oxford, 2000.

[3] P. Bamberg, S. Sternberg, *A course in Mathematics for Students of P hysics: 1*, Cambridge University Press, Cambridge, 2001.

[4] J. R. Cardoso, *Logaritmos de Matrizes – Aspectos Teóricos e Numéricos*, Tese de Doutoramento, Departamento de Matemática, Universidade de Coimbra, 2003.

[5] S. Fiori, Averaging over the Lie group of optical systems transference matrices, *Frontiers of Electrical and Electronic Engineering in China*, 6 (1), pp. 137 – 145, 2011.

[6] F. R. Gantmacher, *Theory of Matrices*, Vol. I e II, Chelsea, New York, 1959.

[7] W. F. Harris, Paraxial ray tracing through noncoaxial astigmatic optical systems, and a 5×5 augmented system matrix, *Optometry and Vision Science*, 71, pp. 282–285, 1994.

[8] W. F. Harris, Diopric power: its nature and its representation in three- and four-dimensional space, *Optometry and Vision Science*, 74, pp. 349–366, 1997.

[9] W. F. Harris, The average eye, *Ophthalmic and Physiological Optics*, 24, pp. 580–585, 2004.

[10] W. F. Harris, Eigenvalues of the transferences of Gaussian optical systems, *South African Optometrist*, 64 (4), pp. 150–153, 2005.

[11] W. F. Harris e J. R. Cardoso, The exponential-mean-log-transference as a possible representation of the optical character of an average eye, *Ophthalmic and Physiological Optics*, 26, pp. 380-383, 2006.

[12] N. J. Higham, *Functions of Matrices – Theory and Computation*, SIAM, Philadelphia, PA, USA, 2008.

[13] R. A. Horn e C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, Paperback Edition, 1994.

[14] M. P. Keating, A system matrix for astigmatic optical systems: I. Introduction and dioptric power relations, *American Journal of Optometry and Physiological Optics*, 58, pp. 810–819, 1981.

[15] W. F. Long, A matrix formalism for decentration problems, *American Journal of Optometry and Physiological Optics*, 53, pp. 27–33, 1976.