
Introdução aos modelos matriciais – A Matriz de Leslie

Quando estudamos o crescimento populacional, quer em termos discretos ($N_t=f(N_{t-1})$) quer em termos contínuos ($dN/dt=f(N)$), não distinguimos entre os indivíduos no que respeita à sua capacidade para sobreviver e para ter filhos. É sabido, contudo, que estas características diferem entre indivíduos com idades diferentes, tamanho corporal diferente, ou quaisquer outras características individuais que influenciem sobrevivência e fertilidade. Uma população formada predominantemente por indivíduos jovens, deve ter características dinâmicas diferentes de uma outra com o mesmo efectivo mas dominada por indivíduos velhos. Assim, uma descrição mais rigorosa do crescimento da população, passa por tomar em consideração que esta se subdivide em grupos de indivíduos com *taxas vitais* (i.e. taxa de sobrevivência e de natalidade) diferentes.

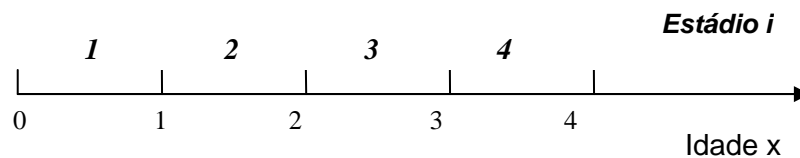
Teóricamente, o critério mais adequado para subdividir os indivíduos, seria aquele que formasse grupos internamente muito homogéneos quanto às taxas vitais e muito distintos entre si quanto às mesmas taxas. Na prática, é muito difícil classificar cada indivíduo com base numa probabilidade *a priori* de sobreviver ou de ter descendentes por unidade de tempo. É mais fácil recorrer a variáveis que se julgam muito correlacionadas com as taxas vitais e que são facilmente mensuráveis nos indivíduos. O exemplo mais comum nas populações humanas é a *idade*. Em populações não-humanas, recorre-se umas vezes à idade outras vezes a medidas do tamanho corporal. Neste texto assume-se que a população se subdivide em grupos de idade ou *grupos etários* de igual duração. A população tem, portanto, aquilo que vulgarmente se designa por uma **estrutura etária**, ou seja, N_x indivíduos na idade x , para $x=0, 1, 2, 3 \dots$. O número total de indivíduos é dado por $N=\sum N_x$. O estudo do crescimento da população que tem em consideração que esta possui estrutura etária é designado por **demografia**.

A instrumentação matemática mais adequada para estudos demográficos, depende da forma como se pretende encarar as duas escalas de tempo envolvidas – o tempo biológico e o tempo de calendário ou de projecção da população (ver abaixo). Há quatro combinações possíveis: considerar ambas as escalas de tempo discretas, ambas contínuas ou a primeira discreta e a segunda contínua ou vice-versa. Neste texto segue-se a primeira opção, i.e., considera-se adequado tratar as duas escalas de tempo de forma discreta. Esta é também a opção mais frequente, provavelmente por razões práticas. Os modelos matemáticos resultantes são do domínio da álgebra matricial, são fáceis de implementar computacionalmente e existe uma vasta panóplia de técnicas de estimação dos valores numéricos das taxas vitais por grupo etário (não abordada aqui) destes modelos.

Os modelos matriciais de populações serão o objecto de estudo deste texto. Estes modelos têm raízes em trabalhos de meados da década de 1940 e devem-se aos esforços de Leslie (1945), mais tarde refinados por Lefkovitch (1965). Actualmente, a principal referência literária para este tipo de modelos é o livro de Caswell (2001).

11.1 Tempo biológico e tempo de projecção.

Considere-se o tempo decorrido entre o nascimento e a morte de um indivíduo típico de determinada espécie biológica em estudo. Divida-se a variável contínua, que representa o tempo de vida, em intervalos de **tempo biológico**, discretos, que se designam por **estádios** biológicos. Se estes estádios tiverem todos a mesma duração, são coincidentes com aquilo que vulgarmente se designa por idades, x . As idades iniciam-se em $x=0$, mas os estádios, que designo por i , iniciam-se em $i=1$, de acordo com o seguinte esquema,



O estágio i ocupa portanto o intervalo de idades $i-1 < x < i$.

[NOTA – embora neste texto se lide apenas com estádios de igual duração, ou seja idades, a consideração da existência de ‘estádios’ permite passar com facilidade para a abordagem de índole mais geral, na qual o tempo biológico é dividido em intervalos de duração que pode ser diferente; permite também ao estudante interessado entrar com mais facilidade no livro de referência de Caswell (2001) sobre modelos matriciais.]

Quando se descreve o crescimento da população, consideram-se instantes de tempo, t , $t+1$, $t+2$, ... nos quais o número de indivíduos presentes na população é contabilizado. O intervalo de tempo $(t, t+1)$ é designado por **intervalo de projecção**¹. É um intervalo do calendário e, para evitar erros, é aconselhável que respeite duas regras. A primeira é que os intervalos de projecção têm sempre a mesma duração. A segunda regra é que a duração do intervalo de projecção não pode ser maior do que a duração do intervalo de idade. A segunda regra faz com que *um indivíduo não possa saltar duas ou mais idades em um só intervalo de projecção*. Por exemplo, se o tempo biológico foi dividido em intervalos de 1 ano, mas a população vai ser projectada de ano-e-meio em ano-e-meio, há organismos que podem saltar 2 idades na passagem de t para $t+1$, enquanto outros saltam só uma idade. A confusão que isto gera pode ser considerável. Quanto à primeira regra, serve para que não seja necessário mudar os valores dos parâmetros de projecção de cada vez que se avança um intervalo de projecção.

¹ *Projecção* é um termo usado em dinâmica populacional e demografia como sinónimo de “prever o futuro sob pressupostos”. Projectar uma população é, portanto, dizer o que vai acontecer à população, caso se verifiquem pressupostos devidamente explicitados.

11.2 Os parâmetros de projecção: P_i e F_i e a matriz de Leslie

Nesta secção definem-se os parâmetros P_i e F_i , que passarei a utilizar para projectar a população. Para facilitar a compreensão destas definições assumam-se que os estádios biológicos têm todos a duração de 1 ano e que o mesmo se passa com a duração dos intervalos de projecção. Ao longo da vida, um organismo transita sucessivamente do estágio i para o estágio $i+1$ e, dentro de cada estágio, pode originar descendentes. Os momentos em que a população é recenseada, entretanto, posicionam-se algures no meio dos estádios biológicos. Definam-se então P_i e F_i da seguinte forma,

P_i é a probabilidade de que um indivíduo do estado i , no instante de census t , sobreviva e esteja no estágio $i+1$ no próximo census $t+1$. Trata-se de uma probabilidade, $0 \leq P_i \leq 1$. No caso de reprodutores sazonais, quando o census da população ocorre logo a seguir à reprodução, os estádios identificam-se com as idades e P_i é simplesmente a taxa de sobrevivência da idade x : S_x .

F_i é o número de descendentes *viáveis* de um indivíduo no estágio i , produzidos durante o intervalo de projecção $(t, t+1)$. “Viáveis” significa que estes descendentes ainda estão vivos no início do intervalo de projecção seguinte, i.e. no instante $t+1$.

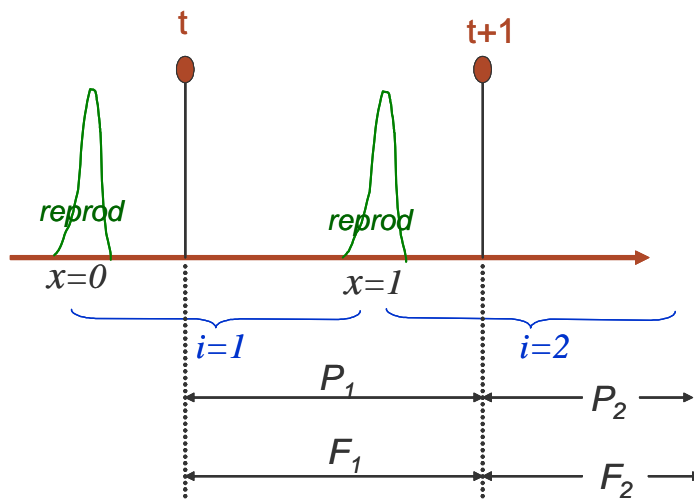


Figura 11.1. Os recém-nascidos iniciam a sua vida na primeira época de reprodução à esquerda (idade $x=0$) e dão início ao estágio $i=1$. A população é contabilizada em t , contando-se $N_{1,t}$ indivíduos no estágio 1 (idade $x=0$). A população volta a ser contabilizada em $t+1$ e, nessa altura, os sobreviventes dos nossos recém-nascidos são $N_{2,t+1}$ e estão no estágio 2 com idade $x=1$. P_1 aplica-se ao intervalo de tempo $(t, t+1)$ e é a probabilidade de sobrevivência dentro deste intervalo. F_1 é a fertilidade e refere-se ao número de filhos nascidos dentro do intervalo $(t, t+1)$ que chegam vivos a $t+1$ e podem ser contabilizados nesse instante.

Quando o census não é feito imediatamente a seguir à reprodução, o *período de tempo* a que P_i e F_i se aplicam é definido pelo intervalo de tempo $(t, t+1)$ entre dois censos consecutivos (Fig 1).

Considere-se a população de indivíduos em t , no início do intervalo de projecção $(t, t+1)$, e designe-se por $N_{1,t}$ o número de indivíduos no estágio 1, $N_{2,t}$ o número de indivíduos no estágio 2, ..., $N_{i,t}$ o número de indivíduos no estágio i . Tendo em atenção que F_i é o número de descendentes *por indivíduo*, no instante $t+1$ o número de indivíduos no estágio 1 é a soma da contribuição de todos os $N_{i,t}$'s:

$$N_{1,t+1} = F_1 N_{1,t} + F_2 N_{2,t} + F_3 N_{3,t} + \dots + F_k N_{k,t} = \sum_{i=1}^k F_i N_{i,t} \quad [11.1]$$

onde k é o último estágio. Os indivíduos no estágio 2, em $t+1$, são os que estavam no estágio 1 e transitaram de estágio,

$$N_{2,t+1} = P_1 N_{1,t}$$

De um modo geral,

$$N_{i+1,t+1} = P_i N_{i,t}, \quad i = 1, \dots, k-1 \quad [11.2]$$

Uma vez definidos os F_i 's e os P_i 's, a projecção da população por [11.1] e [11.2] decorre pacificamente e pode ser repetida sucessivamente para vários intervalos de projecção. Leslie (1945) reconheceu que as operações [11.1] e [11.2] podiam ser representadas por meio de operações de álgebra de matrizes de forma simples. Para isso, é necessário primeiro construir a chamada **matriz de projecção**, **A**, também chamada matriz de transição da população. Os elementos de **A** são os F 's e P 's organizados da seguinte forma:

$$\mathbf{A} = \begin{bmatrix} F_1 & F_2 & F_3 & \dots & F_k \\ P_1 & 0 & 0 & \dots & 0 \\ 0 & P_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P_{k-1} & 0 \end{bmatrix} \quad [11.3]$$

A é uma **matriz quadrada**, por outras palavras, o número de linhas iguala o número de colunas. As contribuições reprodutivas de cada estágio ocupam a primeira linha de **A** e as probabilidades de transição de estágio ocupam a primeira sub-diagonal. Quando tem esta forma (11.3), **A** é designada por **matriz de Leslie**, em reconhecimento do trabalho de Patrick Leslie.

Se a estrutura etária da população no instante t fôr representada pelo vector-coluna **N_t**:

$$\mathbf{N}_t = \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ \dots \\ N_{k,t} \end{bmatrix} \quad [11.4]$$

É fácil verificar que as operações [11.1] e [11.2], para obter a estrutura etária em $t+1$, isto é \mathbf{N}_{t+1} , equivalem a multiplicar \mathbf{N}_t à esquerda por \mathbf{A} :

$$\mathbf{N}_{t+1} = \begin{bmatrix} N_{1,t+1} \\ N_{2,t+1} \\ N_{3,t+1} \\ \dots \\ N_{k,t+1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^k F_i N_{i,t} \\ P_1 N_{1,t} \\ P_2 N_{2,t} \\ \dots \\ P_{k-1} N_{k-1,t} \end{bmatrix} = \begin{bmatrix} F_1 & F_2 & F_3 & \dots & F_k \\ P_1 & 0 & 0 & \dots & 0 \\ 0 & P_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & P_{k-1} & 0 \end{bmatrix} \begin{bmatrix} N_{1,t} \\ N_{2,t} \\ N_{3,t} \\ \dots \\ N_{k,t} \end{bmatrix} \quad [5.12]$$

ou, mais simplesmente,

$$\mathbf{N}_{t+1} = \mathbf{A} \mathbf{N}_t \quad [11.5]$$

Assumindo que os elementos de \mathbf{A} se mantêm constantes no intervalo $(t+1, t+2)$, então a relação entre \mathbf{N}_{t+2} e \mathbf{N}_{t+1} deve ser,

$$\mathbf{N}_{t+2} = \mathbf{A} \mathbf{N}_{t+1}.$$

Mas como $\mathbf{N}_{t+1} = \mathbf{A} \mathbf{N}_t$, então,

$$\mathbf{N}_{t+2} = \mathbf{A} \mathbf{A} \mathbf{N}_t = \mathbf{A}^2 \mathbf{N}_t$$

Assumindo que \mathbf{A} se mantém constante, poder-se-ia obter \mathbf{N}_{t+3} , \mathbf{N}_{t+4} , etc. do mesmo modo. De um modo geral,

$$\mathbf{N}_{t+n} = \mathbf{A}^n \mathbf{N}_t \quad [11.6]$$

A equação [11.6] mostra que a estrutura da população, ao fim de n intervalos de projecção, pode ser obtida calculando a n -ésima potência de \mathbf{A} . Torna-se portanto interessante ter meios para avaliar as características da n -ésima potência de \mathbf{A} . O exemplo numérico a seguir vai permitir deduzir outra propriedade interessante de \mathbf{A} .

11.3 A distribuição etária estável (DEE)

Considere-se a seguinte matriz de projecção com dimensão $(2, 2)$,

$$\mathbf{A} = \begin{bmatrix} F_1 & F_2 \\ P_1 & 0 \end{bmatrix} = \begin{bmatrix} 4.8 & 5.76 \\ 0.242 & 0 \end{bmatrix}$$

Suponhamos que no instante t estavam 240 e 0 indivíduos, respectivamente, nos estádios 1 e 2. A projecção da população para o instante $t+1$, faz-se multiplicando o vector-coluna com a estrutura etária, pela matriz \mathbf{A} (eq. 5.12),

$$\mathbf{N}_{t+1} = \begin{bmatrix} 4.8 & 5.76 \\ 0.242 & 0 \end{bmatrix} \begin{bmatrix} 240 \\ 0 \end{bmatrix} = \begin{bmatrix} 1152 \\ 58.08 \end{bmatrix}$$

onde $1152=(4.8 \times 240)+(5.76 \times 0)$ e $58.08=(0.242 \times 240)+(0 \times 0)$ são, respectivamente, o número de indivíduos no estádio 1 e 2 em $t+1$. Este procedimento pode ser repetido mais vezes, para obter sucessivamente \mathbf{N}_{t+2} , \mathbf{N}_{t+3} , etc. Os resultados seriam,

$$\mathbf{N}_{t+2} = \begin{bmatrix} 5864.14 \\ 278.78 \end{bmatrix} \quad \mathbf{N}_{t+3} = \begin{bmatrix} 29753.64 \\ 1419.12 \end{bmatrix} \quad \mathbf{N}_{t+4} = \begin{bmatrix} 150991.63 \\ 7200.38 \end{bmatrix} \quad \mathbf{N}_{t+5} = \begin{bmatrix} 766234.01 \\ 36539.97 \end{bmatrix}$$

O número absoluto de indivíduos da população, total e por estádio, está a aumentar. Repare-se, contudo, que as *proporções* de cada estádio tendem a estabilizar:

$$\mathbf{N}_{t+1} = \begin{bmatrix} 0.9520 \\ 0.0480 \end{bmatrix} \quad \mathbf{N}_{t+2} = \begin{bmatrix} 0.9545 \\ 0.0454 \end{bmatrix} \quad \mathbf{N}_{t+3} = \begin{bmatrix} 0.9545 \\ 0.0455 \end{bmatrix} \quad \mathbf{N}_{t+4} = \begin{bmatrix} 0.9545 \\ 0.0455 \end{bmatrix} \quad \mathbf{N}_{t+5} = \begin{bmatrix} 0.9545 \\ 0.0455 \end{bmatrix}$$

Este é um resultado fundamental da demografia, deduzido por Alfred Lotka (o mesmo dos modelos Lotka-Volterra de predação e competição). Se a fertilidade por idade e a sobrevivência por idade se mantiverem constantes durante um número suficiente de intervalos de tempo, as proporções de indivíduos em cada idade estabilizam. A população atinge a chamada Distribuição Etária Estável ou DEE.

O quociente entre o número *total* de indivíduos no início dos intervalos (t , $t+1$) sucessivos é, por definição, a taxa de incremento da população, λ_t , que já encontramos quando estudámos o crescimento geométrico desregulado. À medida que a população tende para a DEE, o valor deste quociente tende a estabilizar. Como exercício, podem verificar que o quociente tende para $\lambda=5.075$. Mais interessante, o quociente entre o número sucessivo de indivíduos em *qualquer dos estádios*, também tende a estabilizar no mesmo valor. Considere-se por exemplo o estádio 1. O quociente $N_{1,t+1}/N_{1,t}$ tende a estabilizar, à medida que as projecções avançam. De facto,

$$1152/240=4.8, 5864.14/1152=5.090, \dots, 150991.63/29753.64=5.075, 766234.01/150991.63=5.075$$

Se efectuássemos o mesmo exercício para o estádio 2, calculando sucessivamente $N_{2,t+1}/N_{2,t}$, obteríamos o mesmo resultado. Por outras palavras, *uma vez atingida a DEE*, a população cresce com taxa de incremento, λ , constante. Esta taxa pode ser calculada fazendo o quociente entre o total de indivíduos em t e $t+1$, ou entre o número de indivíduos de qualquer dos estádios. Após a DEE ter sido atingida, é possível portanto obter \mathbf{N}_{t+1} a partir de \mathbf{N}_t simplesmente multiplicando este último por λ .

Neste momento dispomos já da informação necessária para estabelecer uma relação entre λ e \mathbf{A} . Vimos que \mathbf{N}_{t+5} tanto pode ser obtido multiplicando \mathbf{A} por \mathbf{N}_{t+4} como multiplicando λ por \mathbf{N}_{t+4} . Por outras palavras, uma vez estabilizada a estrutura etária, o número (ou escalar, para usar a linguagem da álgebra) λ , produz uma transformação em \mathbf{N} que é equivalente à transformação causada pela matriz \mathbf{A} :

$$\mathbf{A} \mathbf{N} = \lambda \mathbf{N} \quad [11.7]$$

Em álgebra, um escalar com esta propriedade é, por definição, designado por **autovalor** (ou valor próprio) da matriz **A**. Já agora, também por definição, um vector **N** que verifique a igualdade [11.7] designa-se por **autovector** (ou vector próprio) de **A** correspondente a λ . Se **N** é um autovector de **A** que corresponde ao autovalor λ , então $c\mathbf{N}$ também o é, sendo c uma constante qualquer diferente de zero. Uma matriz quadrada de dimensão (n, n) tem, pelo menos, um autovalor diferente de zero e, no máximo, n autovalores diferentes, habitualmente designados por $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$. O *maior* dos autovalores, em termos absolutos, é designado por **raio espectral** da matriz e tem um interesse especial para nós. Vou representá-lo por λ_1 . No exemplo numérico apresentado, $\lambda_1 = 5.075$ é um valor próprio de **A**. Acontece que **A** tem ainda um segundo valor próprio que é $\lambda_2 = -0.275$. O raio espectral é portanto $\lambda_1 = 5.075$.

De um modo geral, é possível demonstrar que *uma vez atingida a DEE, a taxa de incremento da população, λ , é o primeiro autovalor da matriz de projecção da população, λ_1 . Qualquer dos autovectores correspondentes a λ_1 permite determinar a DEE da população.*

Uma vez atingida a DEE, a matriz de projecção pode ser substituída por λ_1 na equação [5.12]:

$$\mathbf{N}_{t+1} = \lambda_1 \mathbf{N}_t \quad [11.8]$$

O número absoluto de indivíduos no futuro, pode ser projectado elevando λ_1 a uma potência igual ao número de intervalos de projecção desde o momento em que a DEE é atingida:

$$\mathbf{N}_t = \lambda_1^t \mathbf{N}_0 \quad [11.9]$$

sendo \mathbf{N}_0 a população inicial em DEE.

Falta só acrescentar que valores e vectores próprios de matrizes quadradas, são entidades perfeitamente respeitáveis em álgebra e que existem técnicas para os calcular directamente a partir de **A**. Como estas técnicas são um pouco morosas, é hábito recorrer a programas computacionais para o fazer.

Literatura Citada

- Caswell, H. 2001 (2nd ed). *Matrix Population Models*. Sinauer, Sunderland, Mass.
- Ebert, T.A. 1999. *Plant and Animal Populations. Methods in Demography*. Academic Press, San Diego, Calif.
- Lefkovich, LP. 1965. The study of population growth in organisms grouped by stages. *Biometrics* **21**:1-18.
- Leslie, PH. 1945. On the use of matrices in certain population mathematics. *Biometrika* **33**:183-212.