Sôbre algumas equações de diferenças relacionadas com a Genética de Populações

FREDERICO PIMENTEL GOMES

Professor Substituto de Matemática da E. S. A.
"Luiz de Queiroz"

ÍNDICE

1 — INTRODUÇÃO

A aplicação das equações de diferenças ao Cálculo de Probabilidades é bem antiga. Mas na Genética de Populações, que é de origem recente, as equações de diferenças têm um papel fundamental, donde a necessidade em que se encontram os geneticistas de estudar os métodos que nos permitem resolvê-las.

Nêste trabalho examinaremos a resolução das equações lineares de diferenças pelo método das funções geratrizes, introduzido por P. S. Laplace (1749-1827) e extensamente utilizado no seu livro "Théorie analytique des probabilités". O método é relativamente simples e elementar, mas pouco conhecido. A sua aplicação se tem mostrado muito cômoda em equações encontradas pelo nosso colega Warwick Estevam Kerr nos seus estudos de Genética de Populações.

Nêste trabalho, que se destina especialmente aos geneticistas, não cogitaremos de dar às demonstrações todo o rigorismo lógico desejável em um trabalho de Matemática pura.

2 — UM PROBLEMA TÍPICO DE GENÉTICA DE POPULAÇÕES

Consideremos uma população onde se encontram indivíduos de constituição genética AA, Aa, aa, com frequências relativas respectivas u_0, v_0, w_0. Se essa população se reproduz por autofecundação, e se u_n, v_n, w_n indicam as frequências relativas na enésima geração, obtemos fàcilmente as equações :

$$u_{n+1} = u_n + \frac{1}{4} v_n , \tag{2,1}$$

$$v_{n+1} = \frac{1}{2} v_n , \tag{2,2}$$

$$w_{n+1} = w_n + \frac{1}{4} v_n . \tag{2,3}$$

Podemos procurar expressões algébricas apenas para u_n e v_n, pois temos

$$u_n + v_n + w_n = 1,$$

logo

$$w_n = 1 - u_n - v_n .$$

A solução de (2,2) é obtida fàcilmente por indução. Não seguiremos, porém, essa marcha, que falha inteiramente nos casos mais complexos.

Troquemos, em (2,1), n por $n + 1$ e obtemos

$$u_{n+2} = u_{n+1} + \frac{1}{4} v_{n+1}$$

$$= u_{n+1} + \frac{1}{4} \cdot \frac{1}{2} v_n ,$$

$$(2,4) \qquad = u_{n+1} + \frac{1}{8} v_n$$

As equações (2,1) e (2,4) nos permitem eliminar v_n e obter a equação final

$$2\, u_{n+2} - 3\, u_{n+1} + u_n = 0,$$

que nos cabe resolver.

3 — UM PROBLEMA TÍPICO DE PROBABILIDADES

Consideremos uma sucessão de infinitas urnas, a primeira com uma bola preta e 5 brancas e cada uma das seguintes com duas pretas e 3 brancas. Um jogador retira uma bola da primeira urna e passa-a à segunda, retira uma da segunda e passa-a à terceira, e assim sucessivamente. Qual será a probabilidade de que seja preta a bola retirada da enésima urna?

Na primeira urna a probabilidade de tirar uma bola preta é $p_0 = \frac{1}{6}$ e a probabilidade de tirar bola branca é $q_0 = \frac{5}{6}$.

Uma vez conhecidas as probabilidades p_n e q_n correspondentes à urna de ordem $n + 1$, para a urna seguinte teremos

$$(3,1) \qquad p_{n+1} = \frac{1}{2} p_n + \frac{1}{3} q_n .$$

pois há duas hipóteses favoráveis:

I) Tirar bola preta na urna de ordem $n + 1$ (p_n) e bola preta na urna seguinte $\left(\frac{2+1}{5+1} = \frac{1}{2}\right)$;

II) Tirar bola branca na urna de ordem $n + 1$ (q_n) e bola preta na urna seguinte $\left(\frac{2}{5+1} = \frac{1}{3}\right)$.

De (3,1) obtemos logo

$$p_{n+1} = \frac{1}{2} p_n + \frac{1}{3} (1 - p_n),$$

$$(3,2) \qquad p_{n+1} = \frac{1}{6} p_n + \frac{1}{3}.$$

Se aí substituirmos n por $n + 1$ obteremos

$$(3,3) \qquad p_{n+2} = \frac{1}{6} p_{n+1} + \frac{1}{3}.$$

A equação (3,2) se pode subtrair de (3,3) e então se chega à equação final

$$(3,4) \qquad 6p_{n+2} - 7p_{n+1} + p_n = 0 ,$$

que nos cabe resolver.

4 — COMO SE APLICA O MÉTODO DAS FUNÇÕES GERATRIZES

Tomemos para exemplo de aplicação a equação

$$(4,1) \qquad 2\,u_{n+2} - 3\,u_{n+1} + u_n = 0,$$

que vimos atrás.

Admitamos a existência de uma função f (x) que, desenvolvida em série, nos dê

$$(4,2) \quad f(x) = u_o + u_1 x + u_2 x^2 + u_3 x^3 + \ldots + u_n x^n + \ldots$$

Teremos então

$$2f(x) = 2u_o + 2\,u_1 x + 2\,u_2 x^2 + 2u_3 x^3 + \ldots,$$

$$-3x\,f(x) = \qquad -3\,u_o x - 3u_1 x^2 - 3u_2 x^3 - \ldots,$$

$$x^2 f(x) = \qquad u_o x^2 + u_1 x^3 + \ldots.$$

A adição das três equações nos dá

$$(x^2 - 3x + 2)\,f(x) = 2u_o + x\,(2u_1 - 3u_o) +$$

$$+ x^2\,(2u_2 - 3u_1 + u_o) + x^3\,(2u_3 - 3u_2 + u_1) +$$

$$+ \ldots + x^{n+2}\,(2u_{n+2} - 3u_{n+1} + u_n) + \ldots$$

Todos os coeficientes do segundo membro, com exceção dos dois primeiros, se anulam de acôrdo com (4,1). E fica

$$f(x) = \frac{2u_o + x\,(2u_1 - 3u_o)}{x^2 - 3x + 2}.$$

Suponhamos que, na população inicial, as frequências relativas eram

$$u_o = 0{,}2 \quad , \quad v_o = 0{,}4 \quad , \quad w_o = 0{,}4 \ .$$

Temos então, de acôrdo com (2,1),

$$u_1 = 0{,}2 + 1/4 \ . \ 0{,}4 = 0{,}3$$

e fica

$$(4{,}3) \qquad f\,(x) = \frac{0{,}4}{x^2 - 3x + 2}$$

Consideremos, de uma maneira mais geral, uma função

$$f\,(x) = \frac{K}{ax^2 + bx + c}$$

e procuremos desenvolvê-la em série. Sendo *r* e *s* as raízes da equação $ax^2 + bx + c = 0$, temos

$$f\,(x) = \frac{K}{a\,(r - x\ (s - x)} = \frac{K}{ars} \cdot \frac{1}{1 - \frac{x}{r}} \cdot \frac{1}{1 - \frac{x}{s}} \ .$$

Mas

$$\frac{1}{1 - z} = 1 + z + z^2 + z^3 + \ldots ,$$

logo temos

$$\frac{1}{1 - \frac{x}{r}} = 1 + \frac{x}{r} + \frac{x^2}{r^2} + \frac{x^3}{r^3} + \ldots ,$$

$$\frac{1}{1 - \frac{x}{s}} = 1 + \frac{x}{s} + \frac{x^2}{s^2} + \frac{x^3}{s^3} + \ldots ,$$

e portanto

$$(4,4)\quad f(x) = \frac{K}{a\,r\,s}\left[1 + x\left(\frac{1}{r} + \frac{1}{s}\right) + x^2\left(\frac{1}{r^2} + \frac{1}{rs} + \frac{1}{s^2}\right) + \ldots + x^n\left(\frac{1}{r^n} + \frac{1}{r^{n-1}s} + \ldots + \frac{1}{s^n}\right) + \ldots\right]$$

A comparação de (4,2) com 4,4) nos mostra logo que

$$u_n = \frac{K}{a\,r\,s}\left(\frac{1}{r^n} + \frac{1}{r^{n-1}s} + \ldots + \frac{1}{s^n}\right).$$

A expressão entre parênteses não é mais que uma soma dos termos de uma progressão geométrica de razão igual a $\frac{r}{s}$. logo temos

$$u_n = \frac{K}{a\,r\,s} \cdot \frac{\frac{1}{r^n} - \frac{1}{s^n} \cdot \frac{r}{s}}{1 - \frac{r}{s}} = K\,\frac{\frac{1}{r^{n+1}} - \frac{1}{s^{n+1}}}{s - r}$$

onde se supõe $1 - \frac{r}{s} \neq 0$, logo $r \neq s$.

No caso da função geratriz de (4,3) temos $K = 0{,}4$, $r = 1$, $s = 2$, e portanto fica

$$u_n = 0{,}4\left(1 - \frac{1}{2^{n+1}}\right).$$

Quando cresce o número de gerações, isto é, quando $n \to \infty$, obtemos

$$\lim_{n \to \infty} u_n = 0{,}4$$

A marcha que acabamos de ver se aplica a qualquer equação linear homogênea de diferenças, isto é, do tipo

$$a_o u_{n+k} + a_1 u_{n+k-1} + a_2 u_{nk-2} + \ldots + a_k u_n = 0.$$

Quando a equação não é homogênea, é geralmente fàcil reduzí-la ao caso anterior.

Por exemplo a equação (3,2), que não era homogênea, nos deu (3,4), que já é homogênea.

Mas há um artifício geralmente mais vantajoso. Com efeito consideremos a equação

$$(4,5) \qquad p_{n+1} = \frac{1}{6} p_n + \frac{1}{3}$$

e tomemos $p_n = y_n + a$, onde a é uma constante a determinar. Temos então $p_{n+1} = y_{n+1} + a$ e (4,5) nos dá

$$y_{n+1} + a = \frac{1}{6} \quad y_n + \frac{1}{6} a + \frac{1}{3}$$

logo

$$y_{n+1} - \frac{1}{6} y_n = \frac{1}{3} - \frac{5}{6} a .$$

Basta impor a condição

$$\frac{1}{3} - \frac{5}{6} a = 0 ,$$

logo $a = \frac{2}{5}$, para obtermos a equação homogênea

$$(4,6) \qquad y_{n+1} - \frac{1}{6} y_n = 0.$$

E' facil verificar por indução que tôda equação do tipo

$$(4,7) \qquad t_{n+1} - b. t_n = 0$$

tem por solução

$$t_n = t_o \cdot b^n$$

Logo a solução de (4,6) é

$$y_n = y_o \cdot \left(\frac{1}{6}\right)^n .$$

E como $p_n = y_n + a = y_n + \frac{2}{5}$, fica

$$p_n = y_o \left(\frac{1}{6}\right)^n + \frac{2}{5}$$

$$= \left(p_o - \frac{2}{5}\right)\left(\frac{1}{6}\right)^n + \frac{2}{5} .$$

$$= \frac{2}{5} - \frac{7}{30}\left(\frac{1}{6}\right)^n .$$

Conclui-se que, se $n \to \infty$, $p_n \to \frac{2}{5}$. Se interessasse apenas êste limite, êle poderia ter sido obtido mais fàcilmente como se segue. Seja

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} p_{n+1}$$

e (3,2) nos dá

$$p = \frac{1}{6} p + \frac{1}{3} \quad , \text{logo} \quad p = \frac{2}{5} .$$

Êste é o limite procurado.

5 — MAIS UM EXEMPLO DE APLICAÇÃO FREQUENTE

Uma equação de diferenças bastante frequente é

$$p_{n+2} = p_n + p_{n+1}$$

Dela obtemos logo

$$f(x) = \frac{(p_o - p_1)\, x - p_o}{x^2 + x - 1} .$$

As raízes do denominador são

$$r = \frac{-1 + \sqrt{5}}{2} , \qquad s = \frac{-1 - \sqrt{5}}{2} .$$

Logo

$$f(x) = (h\,x + j)\ \frac{1}{r - x} \cdot \frac{1}{s - x} ,$$

$$\text{onde} \quad h = p_o - p_1 , \; j = - p_o .$$

Obtemos então

$$f(x) = \frac{(hx + j)}{r\,s} \left[1 + x\left(\frac{1}{r} + \frac{1}{s}\right) + x^2\left(\frac{1}{r^2} + \frac{1}{r\,s} + \frac{1}{s^2}\right) + \dots + x^{n-1}\left(\frac{1}{r^{n-1}} + \frac{1}{r^{n-2}\,s} + \dots + \frac{1}{s^{n-1}}\right) + x^n\left(\frac{1}{r^n} + \frac{1}{r^{n-1}\,s} + \dots + \frac{1}{s^n}\right) + \dots \right]$$

e portanto

$$p_n = \frac{j}{rs}\left(\frac{1}{r^n} + \frac{1}{r^{n-1}s} + \dots + \frac{1}{s^n}\right) +$$

$$+ \frac{h}{rs}\left(\frac{1}{r^{n-1}} + \frac{1}{r^{n-2}s} + \dots + \frac{1}{s^{n-1}}\right)$$

$$= j\,\frac{\dfrac{1}{r^{n+1}} - \dfrac{1}{s^{n+1}}}{s-r} + h\,.\,\frac{\dfrac{1}{r^n} - \dfrac{1}{s^n}}{s-r}$$

Suponhamos $p_0 = 1$, $p_1 = 1$ e obtemos $h = 0$, $j = -1$, logo

$$p_n = \frac{2^{n+1}}{\sqrt{5}}\left[\frac{1}{(-1+\sqrt{5})^{n+1}} - \frac{1}{(-1-\sqrt{5})^{n+1}}\right]$$

$$= \frac{1}{2^{n+1}\sqrt{5}}\left[(1+\sqrt{5})^{n+1} - (1-\sqrt{5})^{n+1}\right]$$

Esta é a expressão do termo geral da sucessão

1, 1, 2, 3, 5, 8, 13, 21, . . . ,

comumente denominada sucessão de Fibonacci, muito frequente na Genética de Populações.

6 — BIBLIOGRAFIA CONSULTADA

1 — USPENSKY, J. V. — Introduction to Mathematical Probability. McGraw-Hill, New York and London. 1937.

2 — KNOPF, Otto — Cálculo de Probabilidades. Labor, Barcelona — Buenos Aires. 1927.

3 — KERR, Warwick E. — Estudos sôbre a Genética de Populações de Himenópteros em Geral e dos Apíneos Sociais em Particular — Piracicaba. 1950.

4 — DAVIS, Harold T. — The Theory of Linear Operators. The Principia Press, Bloomington, Indiana. 1936.

5 — HOGBEN, Lancelot — An Introduction to Mathematical Genetics. Norton & Company, New York. 1946.