

JERZY BURZYŃSKI

SCHEMATYZACJA METODY MALECOT OBLICZANIA WSPÓŁCZYNNIKÓW POKREWIEŃSTWA I INBREDU OSOBNIKÓW DIPLOIDALNYCH

G. Malécot(1), str. 8, definiuje w następujący sposób współczynnik R_{XY} pokrewieństwa między osobnikami X i Y :

Założmy, że osobniki X i Y mają genotypy (a, b) i (c, d) odpowiednio, gdzie każdy z genów a, b, c, d może być A , lub \bar{A} . Przez R_{XY} oznaczamy prawdopodobieństwo, że gen wylosowany z X jest identyczny dzięki pochodzeniu¹ z genem wylosowanym z Y . Oznaczając przez P ($a = c$) prawdopodobieństwo, że geny a i c są identyczne dzięki pochodzeniu, mamy:

$$R_{XY} = \frac{1}{4} \cdot [P(a=c) + P(a=d) + P(b=c) + P(b=d)] \quad (1)$$

Współczynnik inbrodu F_X osobnika X definiuje Malécot jako prawdopodobieństwo, że dwa geny tego samego locusa osobnika X są identyczne dzięki pochodzeniu, a więc, że są oba kopiami tego samego genu przekazanego przez przodka inbredującego X po dwóch różnych ścieżkach. Jeżeli więc osobnik X ma genotyp (a, b) , to:

$$F_X = P(a=b). \quad (2)$$

Zastępując w (1) Y przez X otrzymujemy współczynnik R_{XX} pokrewieństwa między X a nim samym, wyrażonym przez F_X :

$$R_{XX} = \frac{1}{2}(1 + F_X). \quad (3)$$

Oznaczmy przez $X \times Y$ potomka pary X i Y . Jak wykazuje O. Kempthorne [2], str. 75, jest:

$$\text{zaś} \quad F_{X \times Y} = R_{XY}, \quad (4)$$

$$R_{A, B \times C} = \frac{1}{2}(R_{AB} + R_{AC}). \quad (5)$$

¹ Dwa geny nazywamy identycznymi dzięki pochodzeniu, jeżeli są oba kopiami tego samego genu posiadanego przez przodka, lub jeżeli jeden z genów jest kopią drugiego.

$R_{A, B \times C}$ oznacza tu, zgodnie z przyjętą symboliką, współczynnik pokrewieństwa między A i potomkiem pary B i C . Prostymi konsekwencjami wzoru (5) są:

$$R_{A, A \times B} = \frac{1}{2}(R_{AB} + R_{AA}), \quad (6)$$

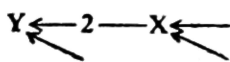
$$R_{A \times B, C \times D} = \frac{1}{4}(R_{AC} + R_{BC} + R_{AD} + R_{BD}), \quad (7)$$

$$R_{A \times B, A \times B} = \frac{1}{4}(R_{AA} + 2R_{AB} + R_{BB}). \quad (8)$$

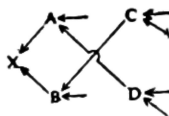
Wzór ostatni przedstawia współczynnik pokrewieństwa między dwoma różnymi bezpośrednimi potomkami pary A i B .

Rodowody osobników, dla których obliczamy R czy F , przedstawiać będziemy w formie strukturalnej. Linie łączące osobniki (oznaczone literami) charakterystyczne dla danego rodowodu zaopatrywać będziemy w strzałki, wskazujące kierunek dziedziczenia genów. Nadto linie zaopatrzone strzałkami (odtąd będziemy pisać krócej: „strzałki”) zaopatrywać będziemy liczbami oznaczającymi odległość pokoleń. Strzałka bez liczby łącząca dwa osobniki oznaczone literami wyrażać będzie, że jeden z nich jest bezpośrednim potomkiem drugiego. Strzałki, których początków nie opatrzono literami, oznaczają przodków nie wpływających na wartość obliczanego współczynnika.

Rodowód strukturalny jest siecią przestrzenną, podobnie jak sieć połączeń radioaparatu, stąd płaski schemat dowolnego rodowodu można wykreślić jedynie przy przyjęciu symbolu przepustu: \leftarrow . Oznacza on, że dwie linie przecinające się na płaszczyźnie rysunku przedstawiają dwie linie w przestrzeni, nie mające wspólnego punktu, tj. osobnika. Rozpatrzmy schematy przykładowe:



Rys. 1

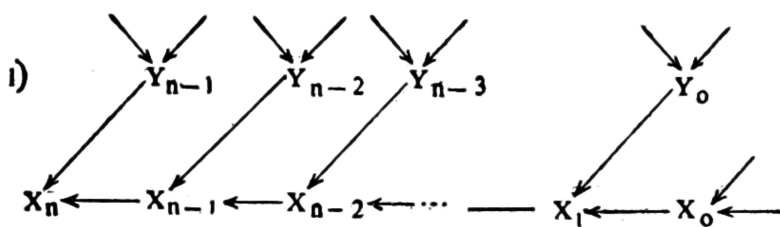


Rys. 2

Z rys. 1 odczytujemy, że Y jest wnukiem, lub wnuczką X . Z rys. 2 wynika, że C i D są rodzicami pełnego rodzeństwa A i B , ci zaś są rodzicami X . Osobnik X jest zimbredowany przez C i przez D .

Nazwijmy węzłem sieci każdy punkt (lub literę zastępującą punkt) przedstawiający: a) osobnika zimbredowanego, b) osobnika inbredującego, c) wspólnego przodka rozpatrywanej pary osobników. Obliczając wartość współczynnika R czy F , będziemy go wyrażać przez współczynniki pokrewieństwa między węzłami najbliższymi punktu wyjściowego (przy obliczaniu F), czy też pary punktów wyjściowych (przy oblicza-

niu R). Postępowanie to będziemy powtarzali aż do natrafienia na osobniki, których współczynnik pokrewieństwa, bądź zinbredowania, jest znany. Podaną wskazówkę objaśnimy przykładami:



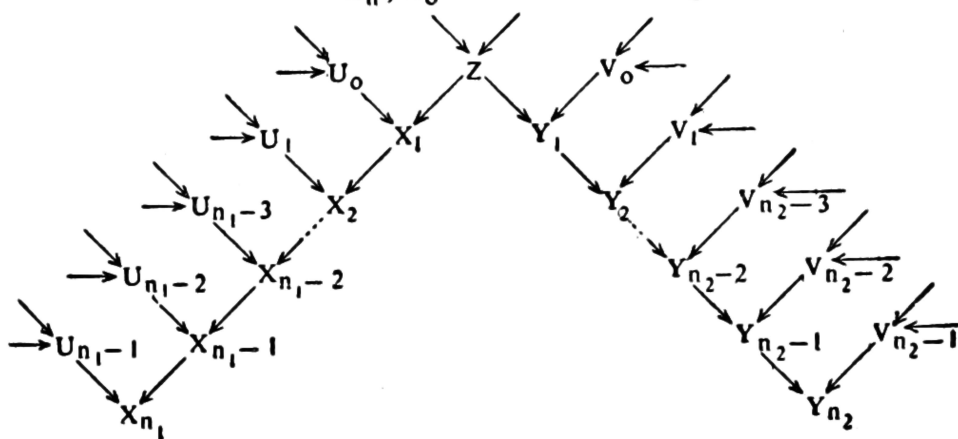
Rys. 3. X_n jest potomkiem X_0 . Żaden z osobników X_1, X_2, \dots, X_n nie jest zinbredowany

Wobec (5) jest

$$R_{X_n, X_0} = R_{X_{n-1} \times Y_{n-1}, X_0} = \frac{1}{2}(R_{X_{n-1}, X_0} + R_{Y_{n-1}, X_0}) = \frac{1}{2}R_{X_{n-1}, X_0}$$

stąd
a wobec (3) $R_{X_n, X_0} = \left(\frac{1}{2}\right)^n \cdot R_{X_0 X_0}$ (9)

$$R_{X_n, X_0} = \left(\frac{1}{2}\right)^{n+1} \cdot (1 + F_{X_0})$$
 (9')



Rys. 4. X_{n1} i Y_{n2} mają jedynego wspólnego przodka Z. Żaden z osobników X_1, X_2, \dots, X_{n1} ani żaden z osobników Y_1, Y_2, \dots, Y_{n2} nie jest zinbredowany.

$$R_{X_{n1}, Y_{n2}} = R_{X_{n1}, Y_{n2-1} \times V_{n2-1}} = \frac{1}{2}(R_{X_{n1}, Y_{n2-1}} + R_{X_{n1}, V_{n2-1}}) = \frac{1}{2}R_{X_{n1}, Y_{n2-1}}$$

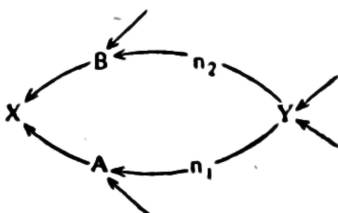
stąd $R_{X_{n1}, Y_{n2}} = \left(\frac{1}{2}\right)^{n2} \cdot R_{X_{n1}, Z}$ (10)

Wobec (9) jest $R_{X_{n1}, Z} = \left(\frac{1}{2}\right)^{n1} \cdot R_{ZZ}$

3

więc ostatecznie $R_{X_{n1}, Y_{n2}} = \left(\frac{1}{2}\right)^{n1 + n2} \cdot R_{ZZ}$ (11)

3)

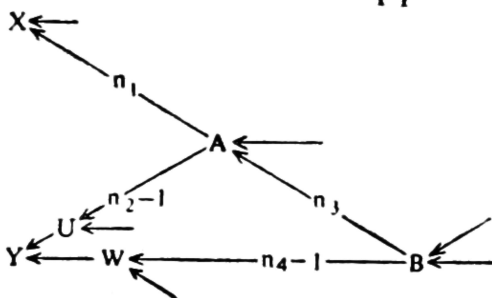


Rys. 5. X jest zinbredowany przez Y. Osobnik Y może być, lub nie być zinbredowany

Wobec (4) i (11) jest

$$F_X = F_{A \times B} = R_{AB} \\ = \left(\frac{1}{2}\right)^{n_1 + n_2} \cdot R_{YY}. \quad (12)$$

4

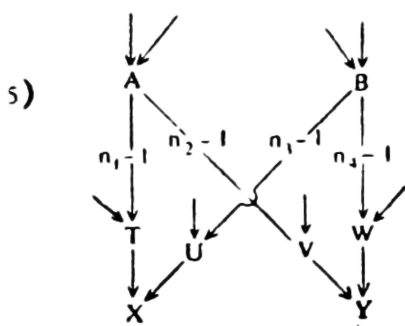


Rys. 6. X i Y mają wspólnego przodka A; również A i Y mają wspólnego przodka B

Na mocy (9) i (10) otrzymujemy kolejno:

$$R_{XY} = \left(\frac{1}{2}\right)^{n_1} \cdot R_{AY} = \left(\frac{1}{2}\right)^{n_1} \cdot R_{A, U \times W} = \left(\frac{1}{2}\right)^{n_1 + 1} \cdot (R_{AU} + R_{AW}) \\ = \left(\frac{1}{2}\right)^{n_1 + 1} \cdot \left[\left(\frac{1}{2}\right)^{n_2 - 1} \cdot R_{AA} + \left(\frac{1}{2}\right)^{n_4 - 1} \cdot R_{AB} \right] \\ = \left(\frac{1}{2}\right)^{n_1 + n_2} \cdot R_{AA} + \left(\frac{1}{2}\right)^{n_1 + n_3 + n_4} \cdot R_{BB}. \quad (13)$$

5

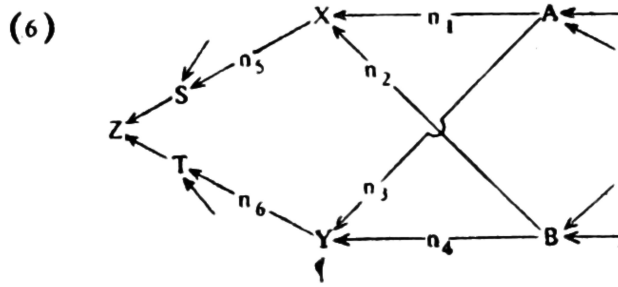


Rys. 7. X i Y mają wspólnych przodków A i B

Stosując (7) i (9) otrzymujemy kolejno:

$$R_{XY} = \frac{1}{4} (R_{TV} + R_{TW} + R_{UV} + R_{UW}) \\ = \frac{1}{4} \left[\left(\frac{1}{2}\right)^{n_1 + n_2 - 2} \cdot R_{AA} + \left(\frac{1}{2}\right)^{n_1 + n_4 - 2} \cdot R_{AB} + \left(\frac{1}{2}\right)^{n_2 + n_3 - 2} \cdot R_{AB} + \left(\frac{1}{2}\right)^{n_3 + n_4 - 2} \cdot R_{BB} \right] \\ = \left(\frac{1}{2}\right)^{n_1 + n_2} \cdot R_{AA} + \left(\frac{1}{2}\right)^{n_3 + n_4} \cdot R_{BB} + \left[\left(\frac{1}{2}\right)^{n_1 + n_4} + \left(\frac{1}{2}\right)^{n_2 + n_3} \right] \cdot R_{AB}. \quad (14)$$

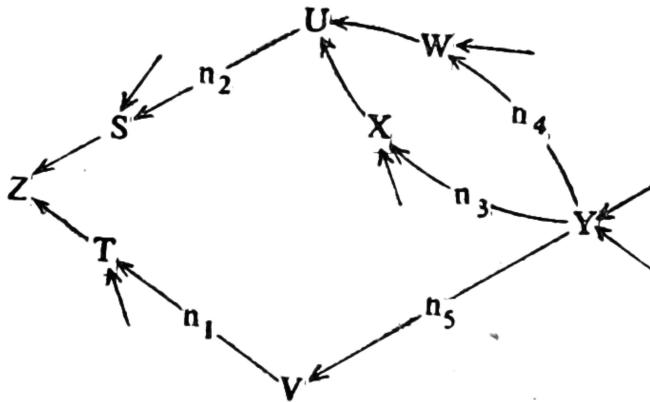
Z wzorów (3) i (14) wynika, że R_{XY} równa się połowie współczynnika korelacji między skojarzonymi zygotami w populacji panmiktycznej, używanego przez S. Wright'a [3] i omówionego szerzej u C. C. Li [4], str. 32 i nast.



Rys. 8. Z jest dwukrotnie zimbredowany, gdyż dwaj jego przodkowie z różnych linii, X i Y, mają dwóch wspólnych przodków A i B

$$F_Z = F_{S \times T} = R_{ST} = \left(\frac{1}{2}\right)^{n_5 + n_6} \cdot R_{XY}, \quad (15)$$

gdzie R_{XY} obliczamy z wzoru (14).



Rys. 9. Z jest dwukrotnie zimbredowany przez tego samego osobnika Y

$$\begin{aligned} F_Z = F_{S \times T} = R_{ST} &= \left(\frac{1}{2}\right)^{n_1 + n_2} \cdot R_{UV} = \left(\frac{1}{2}\right)^{n_1 + n_2} \cdot R_{X \times W, V} \\ &= \left(\frac{1}{2}\right)^{n_1 + n_2 + 1} \cdot (R_{XV} + R_{WV}) \\ &= \left(\frac{1}{2}\right)^{n_1 + n_2 + n_5 + 1} \cdot \left[\left(\frac{1}{2}\right)^{n_3} + \left(\frac{1}{2}\right)^{n_4} \right] \cdot R_{YY}. \end{aligned} \quad (16)$$

Postępując według wyłożonej metody i wykorzystując gotowe wzory tu wyprowadzone, obliczamy bez ryzyka popełnienia błędu i z bardzo małym nakładem pracy wartości R czy F dowolnie skomplikowanych rodowodów. Przy tym zbędne jest wyznaczanie wszystkich ścieżek, po jakich geny identyczne dzięki pochodzeniu mogły zostać przekazane osobnikom spokrewnionym, co stanowi istotę metody S. Wright'a, jak również oryginalnej metody Malécot. Zbędne też okazuje się obliczanie współczynników pokrewieństwa między wszystkimi możliwymi parami osobników objętych rodowodem, co czyni O. Kempthorne w wyżej cytowanym podręczniku.

LITERATURA

1. Malécot G.: Les Mathématiques de l'hérédité; Masson et Cie, Paris 1948.
2. Kempthorne O.: An Introduction to Genetic Statistics. New York, J. Wiley and Sons, Inc., London, Chapman & Hall, Ltd. 1957.
3. Wright S.: Coefficients of inbreeding and relationship. Amer. Nat. 56, s. 330—338.
4. Li C. C.: Population Genetics. The University of Chicago Press.

SCHEMES OF MALÉCOT'S METHOD ON COMPUTING
COEFFICIENTS OF RELATIONSHIP AND INBREEDING AMONG
THE DIPLOID INDIVIDUALS

Malécot defines „coefficient de parenté” R_{XY} as follows: „Consider two individuals X and Y with genotypes ab and cd , where a, b, c , and d may be A or A independently. Then R_{XY} is defined to be probability that a random gene from X is identical by descent with a random gene from Y .”

He defines the coefficient of inbreeding of an individual F_X as the probability that the two genes possessed by individual X at a locus are identical by descent.

Malécot's formulas are applied to find out the mentioned coefficients for lineage of descent in structural forms. In pedigree schemes, besides the individuals which relationship or degree of inbreeding is determined, there are shown only:

(a) inbred individuals, (b) inbreeding individuals, (c) common ancestry of two individuals in question.

The remaining individuals that form the pedigree were left out but for giving the number of generations separating those in (a), (b), (c) from each other.

We simplify the computation by expressing R or F of given two individuals or of an individual by coefficients R and F referring only to common ancestry and to inbreeding individuals.

For typical structures, moreover, the formulas for R_{XY} and F_X depending on the type of structure and on the number of generations separating the individuals characteristic for a given structure, were developed.