

วารสารเกษตร 6,2:128 -134 (2533)

Journal of Agriculture 6,2:128 -134 (1990)

วิธีและเทคนิคการคำนวณด้วยบลัพเพื่อประเมินคุณค่าทางพันธุกรรมของ แม่โคขาวลำพูนสำหรับน้ำหนักหย่านม

สุวัฒน์ รัตนธนาชาติ

BLUP PROCEDURES AND ITS COMPUTING TECHNIQUES TO
DETERMINE GENETIC VALUES OF WHITE LAMPHUN DAMS FOR
WEANING WEIGHT

Suwat Rattanaronchart

ABSTRACT: Computing best linear unbiased predictions of the random elements of the mixed linear models are less simple and time consuming task. Though it was considered to be most efficient method, to determine genetic values of livestock. First use of this method was in dairy cattle. The method now is used also in beef cattle and swine. This paper describes its procedures and computing techniques, using the model to assess dams in the Chiangmai University White Lamphun cattle herd for weaning weight. The described computing techniques, however, can be used for any other mixed linear model as well.

บทคัดย่อ: การคำนวณสำหรับเบสท์ลิเนียร์อันไบแอสพรีดิกชันในมิกซ์โมเดลอาจจะไม่ง่ายนักและใช้เวลาค่อนข้างมาก แต่ก็เป็วิธีที่มีประสิทธิภาพที่สุดสำหรับการประเมินคุณค่าทางพันธุกรรมของปศุสัตว์ ถูกนำมาใช้เป็นครั้งแรกในโคนม และขณะนี้ก็ใช้ในโคเนื้อและสุกรด้วยเหมือนกัน ในบทความนี้ได้อธิบายถึงวิธีและเทคนิคในการคำนวณของวิธีดังกล่าว โดยใช้โมเดลสำหรับการประเมินแม่โคในฝูงโคขาวลำพูนของมหาวิทยาลัยเชียงใหม่สำหรับน้ำหนักหย่านม อย่างไรก็ตาม เทคนิคการคำนวณที่นำเสนอนี้สามารถใช้ได้กับมิกซ์โมเดลทั่วไป

คำนำ

การคัดเลือกแม่โคเนื้อและโคนมในประเทศไทย ส่วนใหญ่จะเน้นกันเฉพาะความสามารถในการให้ลูกเท่านั้น ซึ่งต่างกับประเทศสหรัฐอเมริกาหรือประเทศที่เจริญแล้วอื่นๆ ที่นอกจากจะดูความสามารถในการให้ลูกแล้ว ยังพิจารณาถึงความสามารถของแม่โคในลักษณะอื่นๆ ที่มีความ

สำคัญทางเศรษฐกิจอีกด้วย เช่น น้ำหนักเมื่อหย่านมของลูก และอัตราการเจริญเติบโตของลูก หลังหย่านม เป็นต้น

การคัดเลือกแม่โคสำหรับน้ำหนักเมื่อหย่านมของลูก โดยทั่วไปพิจารณาจากค่าโมสท์ พร็อพแบเบิลโพรคิวซิงอับิลิตี้ หรือที่เรียกโดยย่อว่า เอ็มพีพีเอ (Most Probable Producing Ability : MPPA) แต่ก่อนที่จะทำการคำนวณค่าเอ็มพีพีเอ ข้อมูลจะต้องได้รับการปรับสำหรับความแตกต่างอันเนื่องมาจากฟิക്ซ์แฟคเตอร์ (fixed factor) เช่น เพศของลูก อายุของแม่ ปีและฤดูที่ลูกเกิด เป็นต้น และที่สำคัญอีกประการหนึ่งก็คือ จะต้องทราบค่าอัตราซ้ำ (repeatability) สำหรับน้ำหนักหย่านม ในสหรัฐอเมริกา ค่าที่จะใช้ในการปรับข้อมูลเพื่อให้แม่โคได้รับความยุติธรรมเนื่องจากสภาพแวดล้อม และค่าอัตราซ้ำ นั้นว่ามีพร้อมและค่อนข้างแม่นยำ แต่สำหรับประเทศไทยเรากล่าวได้ว่าไม่มีค่าดังกล่าวในขณะนี้

อย่างไรก็ตาม ด้วยวิธีการของเบสท์ลิเนียร์อันไบแอสพรีดิกชัน หรือที่เรียกโดยย่อว่า บลัฟ (Best Linear Unbiased Prediction: BLUP) ซึ่งนำมาใช้เป็นครั้งแรกในโคนม โดย Henderson (1973) ทำให้ไม่จำเป็นจะต้องปรับข้อมูลก่อนด้วยอิทธิพลของฟิക്ซ์แฟคเตอร์ดังเช่นที่ผ่านมา ถึงกระนั้นก็ตามวิธีการของบลัฟ ก็ยังมีข้อจำกัดอยู่บ้างเกี่ยวกับการคำนวณซึ่งค่อนข้างยุ่งยาก ถึงแม้จะมีคอมพิวเตอร์ช่วยก็ตาม ปัญหาก็คือการขาดโปรแกรมที่สามารถใช้ได้โดยทั่วไป

ดังนั้น วัตถุประสงค์ของบทความนี้ จึงเพื่อที่จะแสดงให้เห็นถึงวิธีและเทคนิคในการคำนวณเพื่อประเมินคุณค่าทางพันธุกรรมของแม่โค ในโครงการอนุรักษ์และปรับปรุงพันธุ์โคขาวลำพูนของมหาวิทยาลัยเชียงใหม่ โดยวิธีบลัฟ และเพื่อเป็นแนวทางสำหรับประยุกต์ใช้กับสัตว์ชนิดอื่นๆ ด้วยเช่นกัน

วิธีการบลัฟ

มิกซ์โมเดลเมื่อคำนึงถึงความเป็นญาติ

มิกซ์โมเดล (mixed model) สำหรับการประเมินคุณค่าทางพันธุกรรมโดยวิธีบลัฟ สำหรับแม่โคในฝูงโคขาวลำพูน มหาวิทยาลัยเชียงใหม่ มีลักษณะดังนี้

$$y = Xb + Z_g u_g + Z_p u_p + e \quad (1)$$

เมื่อ

- y = เวกเตอร์สำหรับข้อมูลน้ำหนักหย่านมปรับเข้าหา 205 วัน และปรับด้วยลิเนียร์เอฟเฟกต์ (linear effect) ของอายุแม่
- b = เวกเตอร์ของฟิക്ซ์เอฟเฟกต์ (fixed effect) ซึ่งประกอบด้วย เพศ และปี-ฤดูที่เกิดของลูกโค
- X = อินซิเดนซ์เมทริกซ์ (incidence matrix) ซึ่งประกอบด้วยอีลีเมนต์ (element) 0 และ 1 สำหรับเชื่อมโยงระหว่างข้อมูลกับอีลีเมนต์ของ b
- u_g = แรนดัมเวกเตอร์ (random vector) ของอิทธิพลแม่โค (transmitting ability effect)

Z_g = อินดิเค้นซ์เมทริกซ์ สำหรับเชื่อมโยงระหว่างข้อมูลกับอีลีเมนต์ของ u_g
 u_p = แรนดัมเวคเตอร์ของเพอร์มาเนนท์เอ็นไวรอนเม้นท์เอฟเฟกต์(permanent environmental effect)
 Z_p = อินดิเค้นซ์เมทริกซ์ สำหรับเชื่อมโยงระหว่างข้อมูลกับอีลีเมนต์ของ u_p
 e = แรนดัมเออเรอะ (random error)

โดยมีสมมุติฐานว่า

$$E \begin{bmatrix} u_g \\ u_p \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{Var} \begin{bmatrix} u_g \\ u_p \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_g^2 & 0 & 0 \\ 0 & I\sigma_p^2 & 0 \\ 0 & 0 & I\sigma^2 \end{bmatrix}$$

เมื่อ

I = ไอเดนทิตีเมทริกซ์ (identity matrix)
 A = เมทริกซ์สำหรับสัมประสิทธิ์ความเป็นญาติของโค (relationship coefficient)
 σ_g^2 = ฟูลสิบโควาเรียนซ์ (full-sib covariance) ดังนั้น
 $\sigma_g^2 = 1/2 h^2 \sigma_t^2$ [h^2 = heritability = 0.3; σ_t^2 = total or phenotypic variance]
 $\sigma_p^2 = (r - 1/2 h^2) \sigma_t^2$ [r = repeatability = 0.4] และ
 $\sigma^2 = (1-r) \sigma_t^2$

สมการมิกซ์โมเดล

บัลสำหรับอีลีเมนต์ของเวคเตอร์ u_g และ u_p สามารถหาได้จากสมการมิกซ์โมเดล (2)

$$\begin{bmatrix} X'X & X'Z_g & X'Z_p \\ Z_g'X & Z_g'Z_g + \frac{\sigma_g^2}{\sigma_g^2} A^{-1} & Z_g'Z_p \\ Z_p'X & Z_p'Z_g & Z_p'Z_p + \frac{\sigma_p^2}{\sigma_p^2} I \end{bmatrix} \begin{bmatrix} b \\ u_g \\ u_p \end{bmatrix} = \begin{bmatrix} X'y \\ Z_g'y \\ Z_p'y \end{bmatrix} \quad (2)$$

โปรดสังเกต ที่จริงแล้วสมการสำหรับ u_g และ u_p คืออันเดียวกัน จะต่างกันก็ที่ $(\sigma^2/\sigma_g^2)A^{-1}$ และ $(\sigma^2/\sigma_p^2)I$ ซึ่งเพิ่มเข้าไปภายหลังเท่านั้น รายละเอียดสามารถศึกษาได้จาก Van Vleck(1979)

การสร้างสมการมิกซ์โมเดล

จะเห็นว่าสิ่งแรกที่จะต้องทำในการหาค่าผลลัพธ์สำหรับ u_g และ u_p ก็คือ การสร้างสมการมิกซ์โมเดล ในกรณีของวิธีการนี้สมการค่อนข้างจะยุ่งยากพอควร เพราะเหตุว่ามีการเพิ่ม $(\sigma^2/\sigma_g^2)A^{-1}$ และ $(\sigma^2/\sigma_p^2)I$ เข้าไปในสมการลีสท์สแควร์ (least-squares equation) ที่คุ้นเคยกัน โดยเฉพาะการเพิ่มของ A^{-1} ทำให้ยังไม่มีผู้ใดเขียนโปรแกรมคอมพิวเตอร์ เพื่อแก้ปัญหาสำหรับมิกซ์โมเดลที่สามารถใช้ได้โดยทั่วไป ส่วนโปรแกรมของ Harvey (1987) สำหรับมิกซ์โมเดลที่ใช้กันแพร่หลายนั้น ใช้ได้เฉพาะการประเมินคุณค่าทางพันธุกรรมสำหรับสัตว์ที่ไม่ได้เป็นญาติกัน และยังมีข้อจำกัดอื่นๆ อยู่อีกบ้าง อย่างไรก็ตาม การหา A^{-1} ที่จริงแล้วง่ายมาก จึงไม่ควรเป็นอุปสรรคแต่อย่างใด (Henderson, 1976)

ก่อนที่จะพูดถึงการสร้างสมการมิกซ์โมเดล จำเป็นต้องทราบถึงปัญหาอีกประการหนึ่ง นั่นก็คือ สมการมิกซ์โมเดลที่ใช้สำหรับการประเมินคุณค่าทางพันธุกรรมของสัตว์ ซึ่งโดยทั่วไปมักมีขนาดใหญ่ ในกรณีของเรานี้ ถึงแม้จะถือว่าเล็กน้อยมาก คือมี 127 สมการก็ตาม แต่ก็นับว่าเป็นปัญหาในการแก้สมการโดยการอินเวอร์ส (inverse) เมทริกซ์ เนื่องจากหน่วยความจำ (RAM) ของเครื่องคอมพิวเตอร์มักจะไม่เพียงพอ ดังนั้นการแก้สมการมิกซ์โมเดลจึงจำเป็นต้องใช้วิธีไอเทอเรชัน (iteration) แทน

เทคนิคการสร้างสมการมิกซ์โมเดลที่เหมาะสมต่อการแก้สมการโดยวิธีไอเทอเรชัน และไม่สิ้นเปลืองหน่วยความจำของเครื่องคอมพิวเตอร์ก็คือ การสร้างสมการมิกซ์โมเดลเก็บไว้ในหน่วยความจำถาวรอย่างใดอย่างหนึ่งของระบบคอมพิวเตอร์ เช่น เทปหรือดิสก์ เป็นต้น และเก็บสมการมิกซ์โมเดลนี้ไว้ภายใต้ชื่อไฟล์ MME.DAT ซึ่งลักษณะของไฟล์ MME.DAT นี้ประกอบด้วย 3 แถวตั้ง (Columns) โดยแถวตั้งที่ 1 ให้ชื่อว่า IROW แถวตั้งที่ 2 ให้ชื่อว่า JCOLUMN และแถวตั้งที่ 3 ให้ชื่อว่า XSUM

เพื่อให้ง่ายและเข้าใจได้เร็วสำหรับการสร้างสมการมิกซ์โมเดลด้วยวิธีดังกล่าว โปรดดูตัวอย่าง พร้อมกับคำอธิบายดังต่อไปนี้ สมมุติว่ามีข้อมูลคือ

เพศ	แม่	น้ำหนักหย่านม
1	3	90
1	4	85
2	3	80
2	4	100
2	5	70

โดยที่แม่หมายเลข 4 เป็นลูกของแม่หมายเลข 3 ดังนั้น ทำให้สามารถเขียนสมการมิกซ์โมเดลได้ดังสมการ (3) ซึ่งรายละเอียดสามารถศึกษาได้จาก Van Vleck (1979)

$$\begin{array}{c}
 1 \\
 2 \\
 3 \\
 4 \\
 5 \\
 6 \\
 7 \\
 8
 \end{array}
 \begin{bmatrix}
 2 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
 0 & 3 & 1 & 1 & 1 & 1 & 1 & 1 \\
 1 & 1 & 2+8 & -4 & 0 & 2 & 0 & 0 \\
 1 & 1 & -4 & 2+8 & 0 & 0 & 2 & 0 \\
 0 & 1 & 0 & 0 & 1+6 & 0 & 0 & 1 \\
 1 & 1 & 2 & 0 & 0 & 2+3 & 0 & 0 \\
 1 & 1 & 0 & 2 & 0 & 0 & 2+3 & 0 \\
 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1+3
 \end{bmatrix}
 \begin{bmatrix}
 b_1 \\
 b_2 \\
 u_{g1} \\
 u_{g2} \\
 u_{g3} \\
 u_{p1} \\
 u_{p2} \\
 u_{p3}
 \end{bmatrix}
 =
 \begin{bmatrix}
 175 \\
 250 \\
 170 \\
 185 \\
 70 \\
 170 \\
 185 \\
 70
 \end{bmatrix}
 \quad (3)$$

ลักษณะของสมการมิชโมเดลเดียวกันนี้ในไฟล์ MME.DAT คือ

IROW	JCOLUMN	XSUM
1	1	2
1	3	1
1	4	1
1	6	1
1	7	1
1	999	175
2	2	3
2	3	1
2	4	1
2	5	1
2	6	1
2	7	1
2	8	1
2	999	250
:	:	:
:	:	:
8	2	1
8	5	1
8	8	1+3
8	999	70

จึงเห็นได้ว่า IROW และ JCOLUMN เป็นตัวระบุตำแหน่งของค่าสำหรับ XSUM ใน (3) หมายเลข 999 ใน JCOLUMN ใช้เพื่อระบุว่าค่า XSUM ในแถวเดียวกันนั้น เป็นค่าทางขวามือของสมการ (3) และโปรดสังเกตว่า จะไม่มีการบันทึกสำหรับอีลิเมนต์ซึ่งเท่ากับ 0 ในไฟล์นี้

การแก้สมการ

ลักษณะของสมการมิกซ์โมเดลในไฟล์ MME.DAT เหมาะต่อการแก้ด้วยวิธีไอเทอเรชันเป็นอย่างยิ่ง ดังจะเห็นได้จากขั้นตอนในการแก้สมการต่อไปนี้

รอบที่ 0 ให้ค่าเริ่มต้น (prior value) สำหรับ b , u_g และ u_p เท่ากับ 0

รอบที่ 1 หา u_{p3} ซึ่งจะได้ว่า $u_{p3} = 70/(1+3) = 17.5$ เสร็จแล้วคำนวณสำหรับตัวอื่นๆ ที่เหลือในสมการ เช่น

$$b_1 = (175 - 1 \times 0 - 1 \times 0 - 1 \times 0 - 1 \times 0)/2 = 87.5$$

$$b_2 = (250 - 1 \times 0 - 1 \times 0 - 1 \times 0 - 1 \times 0 - 1 \times 0 - 1 \times 17.5)/3 = 77.5$$

$$u_{g1} = (170 - 1 \times 87.5 - 1 \times 77.5 + 4 \times 0 - 2 \times 0)/(2+8) = 0.5$$

จนกระทั่งครบหมดทุกตัว จึงเป็นอันสิ้นสุดรอบที่ 1

รอบที่ 2 และต่อไป ค่าวนในทำนองเดียวกันกับรอบที่ 1 เสร็จแล้วตรวจสอบดูว่าค่าสมบูรณ์ (absolute value) ของคำตอบที่เปลี่ยนจากรอบหนึ่งไปยังอีกรอบถัดไป มีตัวใดที่มากกว่า 10^{-6} ถ้าไม่มีเลยก็แสดงว่า ได้คำตอบสุดท้ายแล้ว

เมื่อทราบขั้นตอนการคำนวณและลักษณะของไฟล์ MME.DAT ดังได้กล่าวไปแล้วนั้น การเขียนโปรแกรมคอมพิวเตอร์เพื่อใช้ในการคำนวณก็จะไม่ยาก ถ้าหากผู้เขียนจะนำโปรแกรมฟอร์แทรน (FORTRAN) ที่เขียนมาอธิบายประกอบด้วยในที่นี้ ก็คงจะยืดเยื้อยาวเกินไป และนั้นก็ไม่ใช่จุดประสงค์ของบทความนี้ ดังนั้นจึงไม่กล่าวถึง

คุณค่าทางพันธุกรรมของแม่โค

คุณค่าทางพันธุกรรมสำหรับน้ำนมหนักหย่านมของแม่โค สามารถคำนวณได้ด้วย $u_{gi} + u_{pi}$ โดยวิธีบลัทท์ ทำให้คุณค่าทางพันธุกรรมของแม่โคที่ประเมินได้ สามารถใช้เปรียบเทียบเพื่อการคัดเลือกได้โดยตรง ถึงแม้ว่าจะเป็นสัตว์คนละชั่วอายุ (generation) ก็ไม่ต้องห่วงเรื่องยีนดริฟต์ (genetic trend) เพราะวิธีการนี้ได้ใช้ A^{-1} ในการคำนวณด้วยแล้ว (Henderson, 1975) ส่วนวิธีประเมินด้วยค่าเอ็มพีพีเอ ไม่ได้คำนึงถึงเรื่องยีนดริฟต์แต่ประการใด

วิจารณ์

วิธีบลัทท์ที่ได้เสนอมาทั้งหมดแล้วนั้นเป็นวิธีการที่ไม่ง่ายนัก แต่ก็เป็วิธีที่มีประสิทธิภาพสูงที่สุดที่มีใช้สำหรับการประเมินคุณค่าพันธุกรรมในปัจจุบัน การประเมินคุณค่าพันธุกรรมของสัตว์เลี้ยงอื่นๆ นอกจากโคนม โดยเฉพาะโคเนื้อและสุกร ก็มีแนวโน้มที่จะใช้วิธีบลัทท์มากขึ้น

จะเห็นได้ว่าเทคนิคการคำนวณที่เสนอในบทความนี้ สามารถช่วยตัดปัญหาการอินเวอร์สเมทริกซ์เพื่อแก้สมการที่มีจำนวนมากๆ ลงไปได้ ดังนั้น การคำนวณโดยวิธีบลัทท์จึงไม่น่าจะเป็นปัญหาอีกต่อไป แต่สิ่งที่อาจจะต้องแก้ไขก็คือ ในขั้นตอนของการแก้สมการโดยวิธีไอเทอเรชันขั้นตอนที่เสนอในบทความนี้ถือว่าง่ายที่สุด แต่จากการคำนวณสำหรับ 127 สมการ พบว่าได้ผลลัพธ์สุดท้ายค่อนข้างช้ามาก กล่าวคือต้องคำนวณถึง 202 รอบ สาเหตุประการหนึ่งก็เนื่องมา

จากลักษณะของสมการมิกซ์โมเดล เพราะอีลีเมนต์ในแนวเส้นทแยงมุมมีค่าไม่สูงมากนัก เมื่อเปรียบเทียบกับอีลีเมนต์นอกเส้นทแยงมุม ถึงกระนั้นก็ตาม Quaas and Pollak (1980) ได้เคยกล่าวถึงการเพิ่มประสิทธิภาพสำหรับการแก้สมการลักษณะเช่นนี้โดยวิธีที่เรียกว่า บล็อกไอเทอเรชัน (block iteration) อีกทางหนึ่งที่จะช่วยในการแก้สมการเพื่อให้ได้ผลลัพธ์สุดท้าย เร็วขึ้น ก็คือการลดจำนวนสมการให้น้อยลง โดยการแอบซอร์บ (absorb) สมการสำหรับฟิคซ์เอฟเฟคท์ เข้าไปไว้ในสมการของแรนดัมเอฟเฟคท์ (Searle, 1971)

เทคนิคอันหนึ่งสำหรับการเขียนโปรแกรมเพื่อแก้สมการด้วยวิธีไอเทอเรชัน ที่ควรจะต้องเตรียมไว้ก็คือ ในกรณีที่มีความจำเป็นจะต้องเลิกการคำนวณกลางคัน ก็ควรที่จะสามารถทำได้โดยไม่ต้องไปเริ่มต้นรอบที่ 0 ใหม่ วิธีก็คือจะต้องเก็บผลลัพธ์รอบที่หยุดไว้เป็นค่าเริ่มต้นสำหรับรอบต่อไป

เอกสารอ้างอิง

- Harvey, W. R. (1987). User's Guide for LSMLMW PC-1 Version. Ohio Univ., Columbus, Ohio.
- Henderson, C. R. (1973). Sire evaluation and genetic trend. In Proceedings of the the Animal Breeding and Genetics Symposium in Honor of Jay L. Lush. ASAS and ADSA, Champaign, IL.
- Henderson, C. R. (1975). Use of all relatives in intra-herd prediction of breeding values and producing abilities. J. Dairy Sci. 58 : 1910.
- Henderson, C. R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32 : 69.
- Quaas, R. L. and Pollak, E. J. (1980). Mixed model methodology for farm and ranch beef cattle testing programs. J. Anim. Sci. 51 : 1277.
- Searle, S. R. (1971). Linear Models. John Wiley & Sons, Inc.
- Van Vleck, D. (1979). Notes on the Theory and Application of Selection Principles for the Genetic Improvement of Animals. Cornell Univ., Ithaca, New York.