Az Agrármérnöki MSc szak tananyagfejlesztése
TÁMOP-4.1.2-08/1/A-2009-0010 projekt

ÁLLATGENETIKA

Debreceni Egyetem
Nyugat-magyarországi Egyetem
Pannon Egyetem

A projekt az Európai Unió támogatásával,
az Európai Szociális Alap
társfinanszírozásával valósul meg.

Tenyészértékbecslés, BLUP

***B**est **L**inear **U**nbiased **P**rediction - Variancia Szerkezet*

Lineáris, vagy nem lineáris

Lineáris
Másodrendű polinom

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i}^2 + b_3 X_{3i} + b_4 X_{4i}^2 + b_5 (X_{1i} X_{2i}) + \cdots + \varepsilon_i$$

Nem lineáris

$$Y_i = b_0 e^{-b_1 X_i} \varepsilon_i$$

Log lineáris

$$\ln(Y_i) = \ln(b_0) - b_1 X_i + \ln(\varepsilon_i)$$

Miért lineáris?

A Taylor-sor:

$$Y = f(X)$$

$$Y \approx f(a) + \frac{f^{'}(a)(x-a)}{1!} + \frac{f^{''}(a)(x-a)^2}{2!} + \cdots + \frac{f^{n}(a)(x-a)^n}{n!}$$

$$Y = e^{-X} \qquad Y^{'} = -e^{-X} \qquad Y^{''} = e^{-X}$$

Ha a=0:

Az alacsonyabb rendű együtthatók fontosabbak, mint a magasabb rendűek

Más értékekre is hasonlóan működik, pl. a=0,1:

$$Y = e^{-X}$$

$$Y \approx 1 - x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

$$a = 0{,}1$$

$$Y \approx 1 - x = 1 - 0{,}1 = 0{,}9$$

$$Y = e^{-0{,}1} = 0{,}904837$$

$$Y \approx 1 - x + \frac{x^2}{2!} = 1 - 0{,}1 + \frac{0{,}1^2}{2!} = 0{,}905$$

$$Y \approx 1 - x + \frac{x^2}{2!} + \frac{x^3}{3!} = 1 - 0{,}1 + \frac{0{,}1^2}{2!} + \frac{0{,}1^3}{3!} = 0{,}904833$$

Általánosítva

Bármely, a háttérben lévő ismeretlen függvény közelíthető egy lineáris polinommal (Lineáris Modell).

• Az alacsonyabb rendű együtthatók fontosabbak, mint a magasabb rendűek

• A modell nem szükséges, hogy biológiai függvényen alapuljon

• Még a magas nem-lineáris rendszerek is közelíthetőek alacsonyabb rendű együtthatójú lineáris modellel

•Tisztán leíró

• Lehetővé teszi a modellben szereplő hatásokra vonatkozó hipotézisvizsgálatot

• Lehetővé tesz korlátozott becslést (a kifejtés egy pont körül történik)

Lineáris Modell

- Alkalmas a nem additív genetikai rendszerek (pl. dominancia és episztázis) közelítésére

- A becslés ereje egészen jó, még akkor is, ha a háttérben lévő modell génkölcsönhatásai nem additívak

- A Lineáris Modell alkalmazása széles körben elterjedt az állattenyésztésben

Egy véletlen hatás a Lineáris Modellben

Együtthatók

$$Y_j = b_0 X_{0j} + b_1 X_{1j} + b_2 X_{2j}^2 + b_3 X_{3j} + b_4 X_{4j}^2 + \varepsilon_j$$

Függő változó (tulajdonság)

Független változók

Véletlen hiba

Mátrix elrendezésben

$$Y_1 = b_0 X_{01} + b_1 X_{11} + b_2 X_{21}^2 + b_3 X_{31} + b_4 X_{41}^2 + \varepsilon_1$$

$$Y_2 = b_0 X_{02} + b_1 X_{12} + b_2 X_{22}^2 + b_3 X_{32} + b_4 X_{42}^2 + \varepsilon_1$$

$$\vdots$$

$$Y_n = b_0 X_{0n} + b_1 X_{1n} + b_2 X_{2n}^2 + b_3 X_{3n} + b_4 X_{4n}^2 + \varepsilon_1$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{01} & X_{11} & X_{21}^2 & X_{31} & X_{41}^2 \\ X_{02} & X_{12} & X_{22}^2 & X_{32} & X_{42}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{0n} & X_{1n} & X_{2n}^2 & X_{3n} & X_{4n}^2 \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} + \begin{matrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{matrix}$$

$$Y = XB + \varepsilon$$

Becslés

Legkisebb négyzetek módszere

- Független változók (X)
  - Állandó (fix)
  - Hiba nélkül méri

- Reziduumok (hiba)
  - Véletlen (random)
  - Független, azonos eloszlású, átlaga 0, varianciája σ^2

Független, azonos eloszlású, átlaga 0, varianciája σ^2

$$V(\varepsilon) = E\left[\varepsilon - E(\varepsilon)\right]^2$$

A hiba eloszlása, amelyből minden minta származik, ugyanaz

$$V(\varepsilon) = \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

Nincs környezeti korreláció

$$V(\varepsilon) = I\sigma_e^2$$

Mikor nem lesznek igazak ezek a feltételezések?

A legkisebb négyzetek módszere szerinti becslés (Ordinary Least Squares, OLS)

$$\sum_{j=1}^{n} \varepsilon_i^2 = \varepsilon' \varepsilon = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \qquad \varepsilon_j = Y_j - E(Y_j)$$

$$E(Y_j) = \sum_{i=0}^{k} b_i X_{ij} \qquad \varepsilon_j = Y_j - \sum_{i=0}^{k} b_i X_{ij}$$

$$\varepsilon' \varepsilon = \sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} \left(Y_j - \sum_{i=0}^{k} b_j X_{ij} \right)^2$$

Olyan megoldást kell találni, ahol a reziduumok négyzetösszege a legkisebb.

A legkisebb négyzetek becslései

$$\varepsilon^{'}\varepsilon = \sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} \left(Y_j - \sum_{i=0}^{k} b_j X_{ij} \right)^2$$

$$\frac{\partial(\varepsilon^{'}\varepsilon)}{\partial b_i} = 2\sum_{j=1}^{n} \left(Y_j - \sum_{i=0}^{m} b_i X_{ij} \right) \left[-b_i X_{ij} \right]$$

Az i-t tegyük nullával egyenlővé, és oldjuk meg az egyenletrendszert

A normálegyenletek

$$\begin{bmatrix} \sum x_{0j}^2 & \sum x_{0j}x_{1j} & \cdots & \sum x_{0j}x_{kj} \\ \sum x_{0j}x_{1j} & \sum x_{1j}^2 & \cdots & \sum x_{1j}x_{kj} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_{0j}x_{kj} & \sum x_{1j}x_{kj} & \cdots & \sum x_{kj}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} \sum x_{0j}y_j \\ \sum x_{1j}y_j \\ \vdots \\ \sum x_{kj}y_j \end{bmatrix}$$

$$X'XB = X'Y$$

$$\hat{B} = \left(X'X\right)^{-1}\left(X'Y\right)$$

$$V\left(\hat{B}\right) = \sigma_e^2\left(X'X\right)^{-1}$$

Becslés

$$\hat{Y} = X\hat{B}$$

$$\hat{B} = \left(X'X\right)^{-1}\left(X'Y\right)$$

$$V(\hat{Y}) = V(X\hat{B})$$

$$V(\hat{B}) = \sigma_e^2 (X'X)^{-1}$$

$$V(\hat{Y}) = XV(\hat{B})X'$$

$$V(\hat{Y}) = X(X'X)^{-1}X'\sigma_e^2$$

Általánosított legkisebb négyzetek módszere (Generalized Least Squares, GLS)

- Legkisebb négyzetek módszere
  - Független változók
    - Állandó (fix)
    - Hiba nélkül mérve
  - Reziduumok
    - Véletlen (random)
    - Független, azonos eloszlású, átlaga 0, varianciája σ^2

- Általánosított legkisebb négyzetek módszere
  - Független változók
    - Állandó (fix)
    - Hiba nélkül mérve
  - Reziduumok
    - Véletlen (random)

$$V(\varepsilon) = V$$

Általánosított legkisebb négyzetek módszere

A feladat minimalizálni a $(y-Xb)'V^{-1}(y-Xb)$

A variancia inverzével súlyozva

$$\hat{b}=(X'V^{-1}X)^{-1}(X'V^{-1}y)$$

Ha $V=I\sigma_e^2$

$$\hat{b}=(X'X)^{-1}(X'y)$$

Maximum likelihood (ML) módszer

- Általánosított legkisebb négyzetek módszere
  - Független változók
    - Állandó (fix)
    - Hiba nélkül mérve
  - Reziduumok
    - Véletlen (random)

$$V(\varepsilon) = V$$

- Maximum likelihood
  - Független változók
    - Állandó (fix)
    - Hiba nélkül mérve
  - Reziduumok
    - Véletlen (random)

$$V(\varepsilon) = V$$

$$\varepsilon \approx N(0, V)$$

Maximum likelihood (ML) módszer

$$L = \frac{1}{2\pi^{\frac{N}{2}} \left|V\right|^{\frac{1}{2}}} e^{-\frac{1}{2}\left(y - Xb\right)' V^{-1} \left(y - Xb\right)}$$

A feladat „**b**” maximalizálása

$$\frac{\partial\left(\ln L\right)}{\partial b} = 0$$

$$\ln L = \ln(C) - \frac{1}{2}(y - Xb)' V^{-1} (y - Xb)$$

$$\frac{\partial(\ln L)}{\partial b} = -\frac{1}{2}(y - Xb)'V^{-1}(-X) - \frac{1}{2}(-X)'V^{-1}(y - Xb)$$

$$\frac{\partial(\ln L)}{\partial b} = (y - Xb)'V^{-1}(X)$$

$$(y - Xb)'V^{-1}X = 0$$

$$(y' - (Xb)')V^{-1}X = 0'$$

$$(y' - b'X')V^{-1}X = 0$$

$$b'X'V^{-1}X = y'V^{-1}X$$

$$\hat{b}' = (y'V^{-1}X)(X'V^{-1}X)^{-1}$$

$$\hat{b} = (X'V^{-1}X)^{-1}(X'V^{-1}y)$$

Ugyanaz, mint az általánosított legkisebb négyzetek módszerénél

„**b**” varianciája

$$V(b) = (X'V^{-1}X)^{-1}$$

Ha $V = I\sigma_e^2$

$$V(b) = \sigma_e^2 (X'X)^{-1}$$

Az egyes ismétlések közötti varianciák a generációkkal növekednek

Példaként legyenek itt az alábbi adatok OLS és GLS lineáris regresszióval becsült együtthatói:

y={ **2,28460**, **2,06908**, **1,70690**, **2,02030**, **1,40335**, **1,20441**, **1,22212**, **0,66843**, **1,03762**, **1,06027**};

V={ **0,72 0 0 0 0 0 0 0 0 0**,
0 6,00 0 0 0 0 0 0 0 0,
0 0 12,39 0 0 0 0 0 0 0,
0 0 0 11,30 0 0 0 0 0 0,
0 0 0 0 16,97 0 0 0 0 0,
0 0 0 0 0 19,45 0 0 0 0,
0 0 0 0 0 0 23,45 0 0 0,
0 0 0 0 0 0 0 33,52 0 0,
0 0 0 0 0 0 0 0 40,45 0,
0 0 0 0 0 0 0 0 0 34,46};

X={ **1 0,000**,
1 0,157,
1 0.281,
1 0,333,
1 0.396,
1 0,459,
1 0.491,
1 0,584,
1 0,624,
1 0,657};

Hasonlítsuk össze a becsléseket és a standard hibákat (SE)

	B		VB	
	2,39		0,023	-0,047
OLS	-2,32		-0,047	**0,118**

$$t = \frac{-2,32}{0,118} = 19,46^{**}$$

	B		VB	
	2,30		0,035	-0,083
GLS	-2,03		-0,083	**0,823**

$$t = \frac{-2,03}{0,832} = 2,43^{NS}$$

Következtetések

- Amennyiben a megfelelő hiba struktúrát nem használjuk
  - Téves következtetések fordulhatnak elő
  - Az ismételt sorok miatt drift előfordulhat
- Megoldás
  - Ismételt sorok
  - Vegyük figyelembe a drift okozta varianciát a modellben
  - A drift arányos „F”-el és az additív genetikai varianciával
  - A BLUP részleges megoldást nyújt erre a problémára

***B**est **L**inear **U**nbiased*
***P**rediction - Becslés*

Az OLS hibái független, azonos eloszlásúsak, átlaguk 0, varianciájuk σ^2

$$V(\varepsilon) = E\left[\varepsilon - E(\varepsilon)\right]^2$$

A hiba eloszlása, amelyből minden minta származik, ugyanaz

$$V(\varepsilon) = \begin{bmatrix} \sigma_e^2 & 0 & 0 & 0 \\ 0 & \sigma_e^2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_e^2 \end{bmatrix}$$

A beltenyésztettség ezt megváltoztatja

Rokon egyedek esetében az értékek nullától eltérőek lesznek

$$V(\varepsilon) = I\sigma_e^2$$

A reziduumok függetlenek

Megoldások

- GLS
  - Az adatok varianciáinak és korrelációinak megváltoztatásával oldja meg a problémát

- Mi történik az állandó (fix) hatásokkal?
  - Hogyan lehet korrigálni
    - A környezeti trendekre ellenőrzés nélkül
    - A tenyészet hatására
    - Az év hatására
    - Az alom hatására

Az adatok torzítása

- Tenyészet hatás
  - Kiegyensúlyozott elrendezés esetében nincs probléma
  - A mintában minden családból kell szerepelnie minden tenyészetben
  - A régi megoldás a tenyészeten belüli különbségek értékelése
  - Mi történik akkor, ha a jobb tenyészetek genetikai háttere jobb
- Az állandó (fix) hatásokat korrigálni kell a genetikai különbségekre
- A véletlen (random) hatásokat korrigálni kell az állandó (fix) hatásokra

Az állandó és véletlen hatások szimultán kiegyenlítése

- A független változók elkülönítése
  - Állandó Xb
  - Véletlen Zu

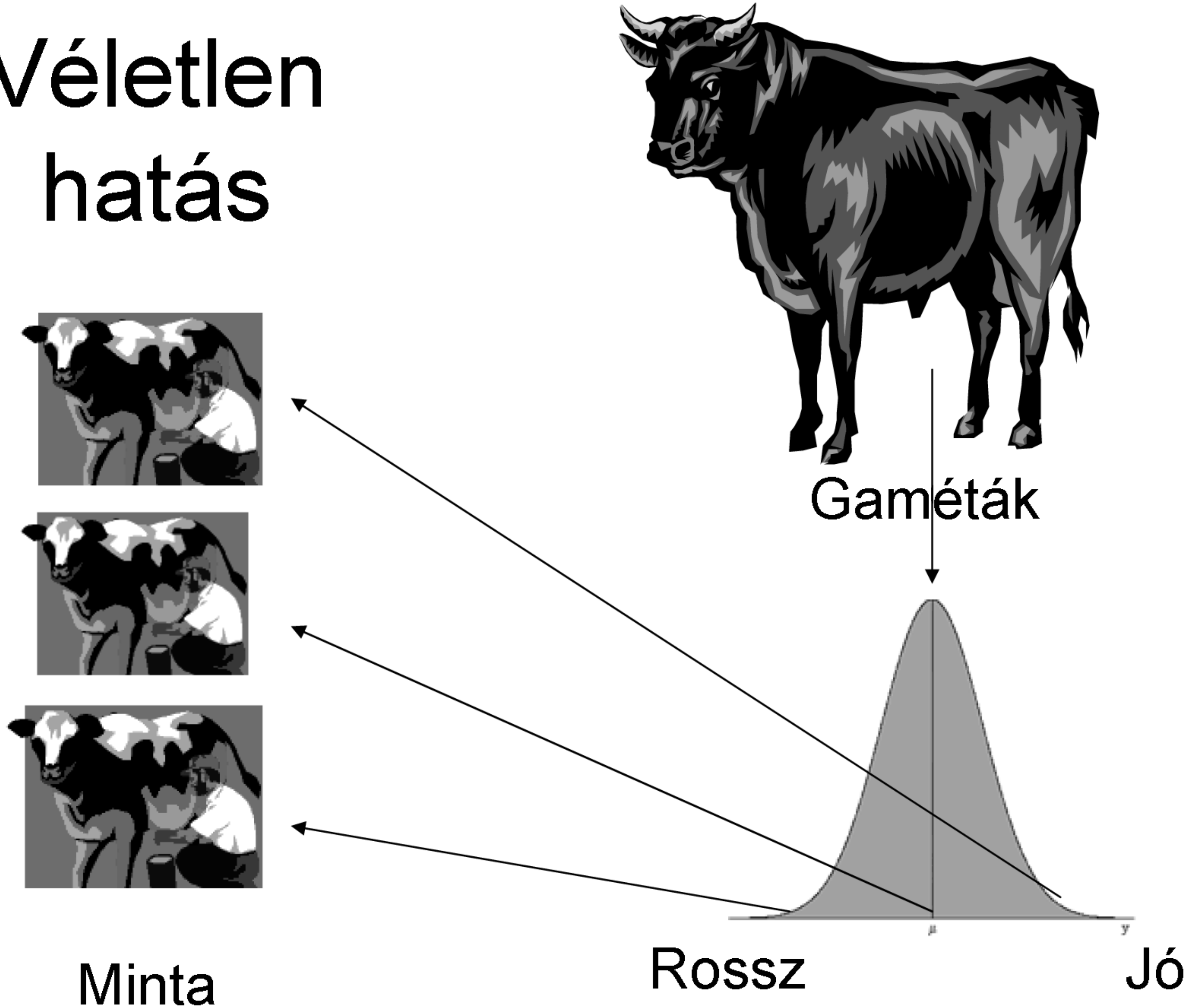
$$Y = Xb + Zu + e$$

Állandó és véletlen hatások

- Állandó hatás
  - A paramétertérben csak a megadott szinteken fejti ki hatását
  - Tenyészet, Év, Évszak, Ellések száma, és Ivar hatása

- Véletlen hatás
  - A hatás a hatások egy eloszlásából vett mintának felel meg
  - A paramétertérben abban a populációban fejti ki hatását, amelyből a véletlen hatást mintázták

Véletlen hatás



Megmutatja a bika genetikai értékét

A vegyes (mixed) modell varianciái

$$Y = Xb + Zu + e \qquad V(b) = 0$$

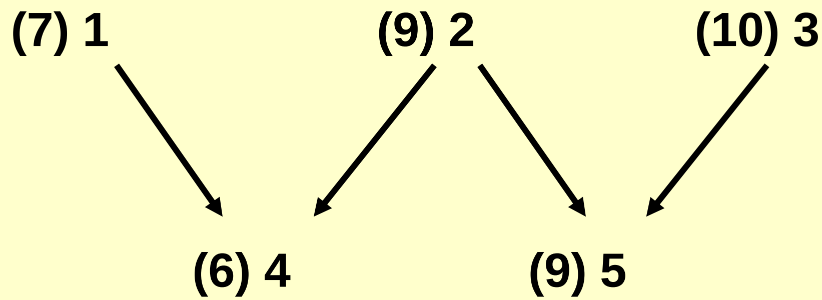
$$V(u) = E(uu') = G$$

$$V(e) = E(ee') = R$$

$$V(Y) = V(Xb + Zu + e) = ZGZ' + R$$

Példa

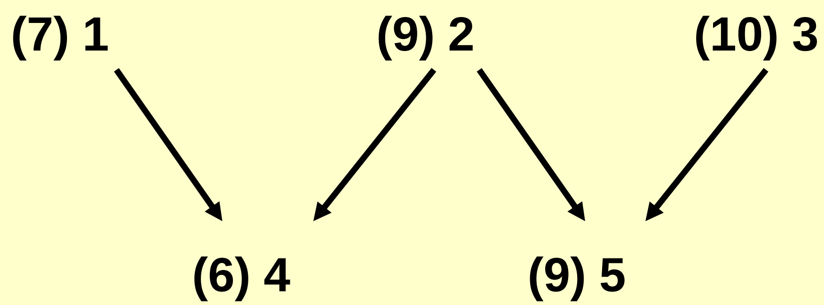
$$Y = \begin{bmatrix} 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix}$$



$$b = [\mu]$$

$$X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad U = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

Példa



$$\begin{bmatrix} 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [\mu] + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

A BLUP ML deriváltja

y és **u** együttes sűrűségfüggvénye: $f(y,u)=g(y/u)h(u)$

$$g(y/u)=g(e)$$

$$g(e)=\frac{1}{(2\pi)^{\frac{1}{2}N}V(e)^{\frac{1}{2}}}e^{-\frac{1}{2}e'V(e)^{-1}e} \qquad h(u)=\frac{1}{(2\pi)^{\frac{1}{2}N}V(u)^{\frac{1}{2}}}e^{-\frac{1}{2}u'V(u)^{-1}u}$$

$$f(y,u)=\frac{1}{(2\pi)^{\frac{1}{2}N}R^{\frac{1}{2}}}e^{-\frac{1}{2}e'R^{-1}e}\frac{1}{(2\pi)^{\frac{1}{2}N}G^{\frac{1}{2}}}e^{-\frac{1}{2}u'G^{-1}u}$$

$$f(y,u)=c_1e^{-\frac{1}{2}e'R^{-1}e}c_2e^{-\frac{1}{2}u'G^{-1}u}$$

$$f(y,u)=ce^{-\frac{1}{2}e'R^{-1}e}e^{-\frac{1}{2}u'G^{-1}u}$$

$$f(y,u) = L = ce^{-\frac{1}{2}e'R^{-1}e} e^{-\frac{1}{2}u'G^{-1}u}$$

A feladat **b** és **u** maximalizálása

$$\ln(L) = \ln(c) - \frac{1}{2}e'R^{-1}e - \frac{1}{2}u'G^{-1}u$$

$$e = Y - Xb - Zu$$

$$\ln(L) = \ln(c) - \frac{1}{2}(Y - Xb - Zu)R^{-1}(Y - Xb - Zu) - \frac{1}{2}u'G^{-1}u$$

Vegyük a **b** deriváltját

$$(Y - Xb - Zu)' R^{-1}(Y - Xb - Zu) + u' G^{-1} u$$

$$= \left[Y' - (Xb)' - (Zu)'\right] R^{-1}(Y - Xb - Zu) + u' G^{-1} u$$

$$= Y' R^{-1} Y - Y' R^{-1} Xb - Y' R^{-1} Zu$$

$$- (Xb)' R^{-1} Y + (Xb)' R^{-1} Xb + (Xb)' R^{-1} Zu$$

$$- (Zu)' R^{-1} Y + (Zu)' R^{-1} Xb + (Zu)' R^{-1} Zu + u' G^{-1} u$$

$$\frac{\partial(\ln L)}{\partial b} = 0 \qquad \begin{aligned} & - Y' R^{-1} X - X' R^{-1} Y + (Xb)' R^{-1} X + \\ & X' R^{-1} Xb + X' R^{-1} Zu + (Zu)' R^{-1} X = 0 \end{aligned}$$

$$- 2X' R^{-1} Y + 2X' R^{-1} Xb + 2XR^{-1} Zu = 0$$

$$X' R^{-1} Xb + X' R^{-1} Zu = X' R^{-1} Y$$

Vegyük az **u** deriváltját

$$(Y - Xb - Zu)' R^{-1} (Y - Xb - Zu) + u' G^{-1} u$$

$$= \left[Y' - (Xb)' - (Zu)' \right] R^{-1} (Y - Xb - Zu) + u' G^{-1} u$$

$$= Y' R^{-1} Y - Y' R^{-1} Xb - Y' R^{-1} Zu$$

$$- (Xb)' R^{-1} Y + (Xb)' R^{-1} Xb + (Xb)' R^{-1} Zu$$

$$- (Zu)' R^{-1} Y + (Zu)' R^{-1} Xb + (Zu)' R^{-1} Zu + u' G^{-1} u$$

$$\frac{\partial (\ln L)}{\partial b} = 0 \qquad \begin{aligned} & - Y' R^{-1} X - X' R^{-1} Y + (Xb)' R^{-1} X + \\ & X' R^{-1} Xb + X' R^{-1} Zu + (Zu)' R^{-1} X = 0 \end{aligned}$$

$$- 2 X' R^{-1} Y + 2 X' R^{-1} Xb + 2 X R^{-1} Zu = 0$$

$$X' R^{-1} Xb + X' R^{-1} Zu = X' R^{-1} Y$$

Vegyük az **u** deriváltját

$$(Y - Xb - Zu)' R^{-1}(Y - Xb - Zu) + u' G^{-1} u$$

$$= Y' R^{-1} Y - Y' R^{-1} Xb - Y' R^{-1} Zu$$

$$- (Xb)' R^{-1} Y + (Xb)' R^{-1} Xb + (Xb)' R^{-1} Zu$$

$$- (Zu)' R^{-1} Y + (Zu)' R^{-1} Xb + (Zu)' R^{-1} Zu + u' G^{-1} u$$

$$\frac{\partial(\ln L)}{\partial u} = 0 \quad \begin{aligned} & -Y' R^{-1} Z + (Xb)' R^{-1} Z - (Z)' R^{-1} Y + (Z)' R^{-1} Xb \\ & + (Z)' R^{-1} Zu + (Zu)' R^{-1} Z + 2G^{-1} u = 0 \end{aligned}$$

$$-2Y' R^{-1} Z + 2Z' R^{-1} Xb + 2Z' R^{-1} Zu + 2G^{-1} u = 0$$

$$Z' R^{-1} Xb + Z' R^{-1} Zu + G^{-1} u = Y' R^{-1} Z$$

A vegyes modell egyenletei

$$X'R^{-1}Xb + X'R^{-1}Zu = X'R^{-1}Y$$

$$Z'R^{-1}Xb + Z'R^{-1}Zu + G^{-1}u = Y'R^{-1}Z$$

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Y'R^{-1}Z \end{bmatrix}$$

Egyszerűsíthető, ha $R = I\sigma_e^2$

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \sigma_e^2 G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

Az „Y” (a függő változó/tulajdonság) normál eloszlása nem szükséges

Alternatív BLUE becslési technikákkal bizonyítható, hogy ugyanezeket az eredményeket kapjuk a normalitás feltételezése nélkül is

Egyszerűsítések

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z+\sigma_e^2 G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

Az additivitást feltételezve: $G = A\sigma_a^2$

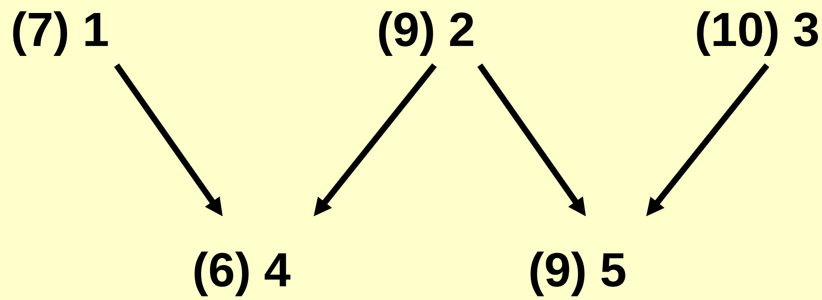
$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z+\frac{\sigma_e^2}{\sigma_a^2} A^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

Csak a hányados becslése szükséges

Csak az inverze szükséges

Példa

$$Y = \begin{bmatrix} 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix}$$



$$b = [\mu]$$

$$X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad U = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \frac{\sigma_e^2}{\sigma_a^2} A^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad X'Z = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

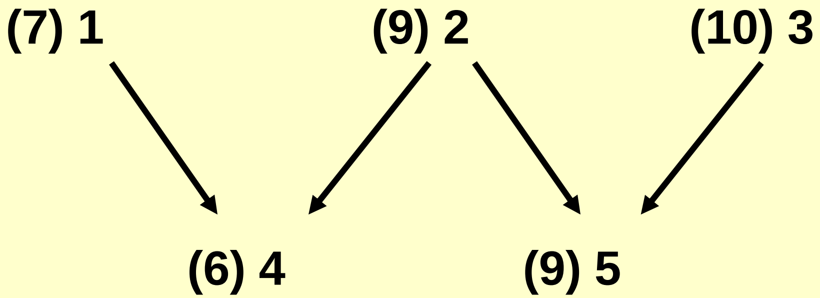
$$X'X = 5 \qquad X'Z = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \dfrac{\sigma_e^2}{\sigma_a^2} A^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

$$Z'X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$Z'Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Z'X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$



$$A = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & 1 \end{bmatrix}$$

Tegyük fel, hogy az örökölhetőség 0,5

$$\frac{\sigma_e^2}{\sigma_a^2} = 1$$

$$Z'Z + \frac{\sigma_e^2}{\sigma_a^2} A^{-1} = \begin{bmatrix} 1 & \frac{1}{2} & 0 & -1 & 0 \\ \frac{1}{2} & 3 & \frac{1}{2} & -1 & -1 \\ 0 & \frac{1}{2} & \frac{5}{2} & 0 & -1 \\ -1 & -1 & 0 & 3 & 0 \\ 0 & -1 & -1 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \frac{\sigma_e^2}{\sigma_a^2} A^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Y \\ Y'Z \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix} \qquad Y'Z = \begin{bmatrix} 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 41 \end{bmatrix}$$

A vegyes modell egyenlete

$$\begin{bmatrix} 5 & 1 & 1 & 1 & 1 & 1 \\ 1 & 5/2 & 1/2 & 0 & -1 & 0 \\ 1 & 1/2 & 3 & 1/2 & -1 & -1 \\ 1 & 0 & 1/2 & 5/2 & 0 & -1 \\ 1 & -1 & -1 & 0 & 3 & 0 \\ 1 & 0 & -1 & -1 & 0 & 3 \end{bmatrix} \begin{bmatrix} \mu \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} 41 \\ 7 \\ 9 \\ 10 \\ 6 \\ 9 \end{bmatrix}$$

A hibavariancia becslése, ha az arány ismert

$$\lambda = \frac{\sigma_e^2}{\sigma_a^2}$$

$$\hat{\sigma}_e^2 = MSE = \left[\frac{Y'Y - \hat{b}X - \hat{u}Z}{N - R(X)} \right]$$

$$\hat{\sigma}_a^2 = \frac{\hat{\sigma}_e^2}{\lambda}$$

A becslések

$$b = \left[\hat{\mu}\right]$$

BU

8,3018868

-0,960813

0,0754717

0,8853411

-1,062409

0,5529753

$$U = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{bmatrix}$$

A becslések varianciái

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \frac{\sigma_e^2}{\sigma_a^2} A^{-1} \end{bmatrix}^{-1}$$

$$V(\hat{b}) = C_{11}\sigma_e^2$$

$$V(\hat{u} - u) = C_{22}\sigma_e^2$$

A becslés hibájának varianciája

$$V(\hat{u}) = A\sigma_a^2 + C_{22}\sigma_e^2$$

A becslés hibájának varianciája,
a drift okozta varianciával együtt

Kennedy és Sorensen, *Quantitative Genetics*

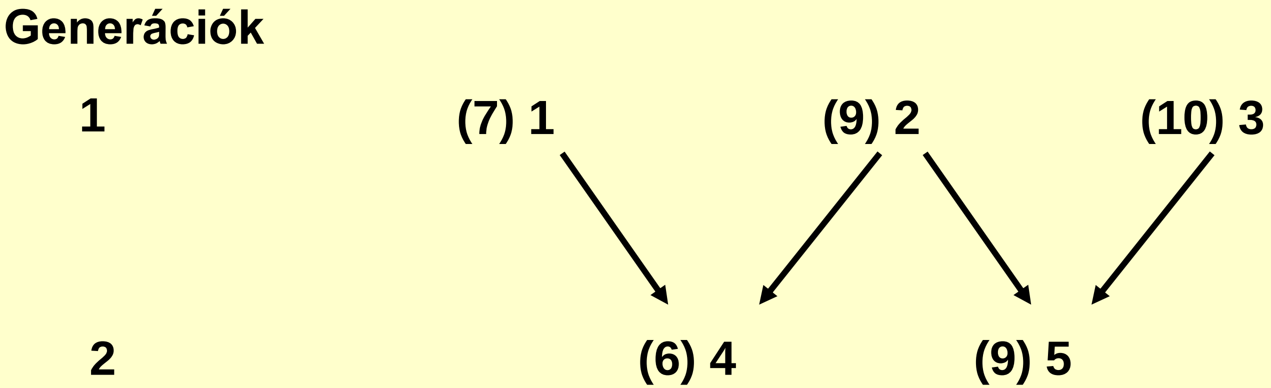
PEV

1,12236	0,29509	0,32030	0,65827	0,39093
0,29509	1,14758	0,29509	0,68854	0,68854
0,32030	0,29509	1,12236	0,39093	0,65827
0,65827	0,68854	0,39093	1,2686	0,60026
0,39093	0,68854	0,65827	0,60026	1,26860

EV

2,86014	0,29509	0,32030	1,52716	0,39093
0,29509	2,88536	0,29509	1,55740	1,55740
0,32030	0,29509	2,86014	0,39093	1,52716
1,52716	1,55740	0,39093	3,00643	1,03471
0,39093	1,55740	1,52716	1,03471	3,00643

Példa

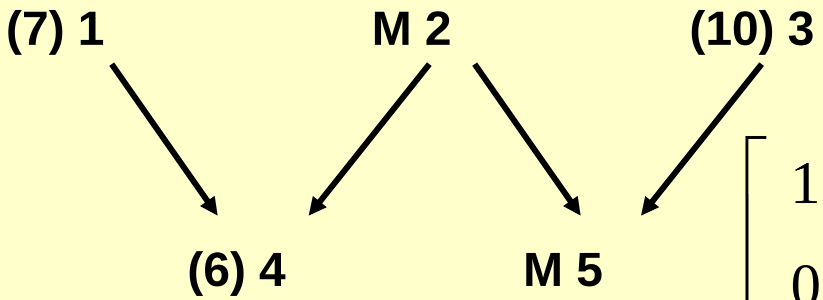


Generáció	Átlag	Szórás
1	8,66	2,860
2	7,5	3,006

Hiányzó értékek
(Ivarhoz kötött tulajdonságok)

Generációk

1



2

$$A=\begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & 1 \end{bmatrix}$$

$$Y=\begin{bmatrix} 7 \\ 10 \\ 6 \end{bmatrix} \qquad e=\begin{bmatrix} e_1 \\ e_3 \\ e_5 \end{bmatrix} \qquad b=\left[\mu\right]$$

$$X=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \; Z=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad U=\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}$$

A becslések

$$b = [\hat{\mu}]$$

7,6153846

-0,307692

-0,589744

0,8974359

-0,448718

-0,435897

$$U = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \end{bmatrix}$$

A modell kiterjesztése

- A dominancia és az episztázis figyelembe vétele
  - A dominancia kapcsoltságának ismerete szükséges
    - Tükröznie kell annak a valószínűségét, hogy az egyedeknek azonos allélpárja van
    - Az episztatikus génhatások additív és domináns génhatások kölcsönhatásainak eredményei
  - Keresztezési programok esetében hasznos a meghatározása, de fajtatiszta tenyésztési programok esetében rendszerint nem szükséges
    - Az egyed nem ad tovább domináns, vagy episztatikus hatást, ez a két szülő eredménye

Limitáló tényezők

- Kiindulópontja, hogy a tulajdonságokat végtelen sok allél befolyásolja
  - Nem működik olyan tulajdonságokra, amelyeket kevés gén befolyásol
  - A genetikai varianciát állandónak feltételezi, kivéve a Bulmer effektust (a variancia csökkenése, mivel nincs egyensúly)
- A modellnek helyesnek kell lennie
  - Hibás bemenő adat esetén az eredmény is hibás lesz
  - A tipikus egyedmodell (Animal Model) feltételezi a rezidiuumok additivitását és függetlenségét

Előadás összefoglalása

- A Lineáris Modell
- A legkisebb négyzetek (OLS), és az általánosított legkisebb négyzetek (GLS) módszere
- A maximum likelihood (ML) becslés
- A BLUP modell egyenlete
- Állandó (fix) és véletlen (random) hatások

Előadás ellenőrző kérdései

- A Lineáris Modell általános felépítése
- Mi alapján dönti el, hogy egy hatás állandó (fix), vagy véletlen (random)?
- Hogyan veszi figyelembe a BLUP a származási adatokat?

KÖSZÖNÖM FIGYELMÜKET

Következő
ELŐADÁS/GYAKORLAT CÍME
Anyai hatás

- **Előadás anyagát készítették: Dr. Posta János és Dr. Komlósi István, Dr. Bruce Walsh és Dr. Michael Lynch tananyagának adaptációjával**