

# Genetika ve šlechtění zvířat

část 9. (rough draft version)

## 10 Smíšené modely v genetických analýzách

Aplikace smíšených modelů je v současné době rozšířený nástroj pro ohodnocování zvířat ve šlechtitelských programech šlechtitelských organizací. Metodologie obsahuje soustavu se statistickými a genetickými vlastnostmi a odvozuje nejpřesnější a nejméně vychýlené předpovědi plemenných hodnot. Kvalita vyhodnocení závisí na:

- datech (záznamy kontroly užitekosti, správná identifikace, správný původ),
- modelu.

Metodologie BLUP má tu vlastnost, že objasňuje selekci rodičů ve šlechtěné populaci. Správně analyzuje skutečnost, že určitá zvířata pocházejí od lepších rodičů, než jiná. Informace o rodokmenu jsou tedy významné a data o vybraných rodičích, stejně jako neselektovaných vrstevníků, jsou začleněny do analýzy. Modely mohou být rozšířené, aby mohly analyzovat komplikované efekty, jako jsou:

- různá plemena (výhodné pro porovnání napříč plemeny, nebo jedince importované z cizích zemí),
- maternální efekty (významné pro všechny vlastnosti před odstavením),
- korelované vlastnosti (pro vyšší přesnost nebo vyhodnocení selekce na druhou vlastnost),
- interakce mezi genotypem a prostředím (určení otcové mohou mít rozdílný efekt v různých prostředích),
- heterogenní variance (rozdíly ve stádě mohou být průměrně větší než rozdíl v jiném stádě).

Určité faktory jsou více obtížné, aby se mohly začlenit do modelu.

### Obecná forma smíšeného modelu:

Model:  $\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$

kde  $\mathbf{b}$  je vektor pevných efektů s designovou maticí  $\mathbf{X}$   
 $\mathbf{u}$  je vektor náhodných efektů s designovou maticí  $\mathbf{Z}$

Definice modelu  $E(\mathbf{y}) = \mathbf{Xb}$   
 $\text{var}(\mathbf{u}) = \mathbf{G}$   
 $\text{var}(\mathbf{e}) = \mathbf{R}$   
 $\text{var}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$

Normální rovnice

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{H} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

Tato základní struktura se rozšiřuje mnoha způsoby. Vektor  $\mathbf{u}$  by mohl obsahovat více náhodných efektů (např. aditivně genetické, maternálně genetické, permanentní prostředí, maternální prostředí, ...). Efekt  $\mathbf{u}$  je určován strukturou matice  $\mathbf{G}$ . Vektor  $\mathbf{y}$  by mohl obsahovat také více vlastností a tedy vektor  $\mathbf{u}$  by mohl mít plemenné hodnoty vztažené k více

vlastnostem zároveň. Matice  $\mathbf{R}$  by mohla obsahovat korelace mezi rezidui (chybami), tj. s korelovanými vlastnostmi. Je třeba v modelu definovat nejen pevné efekty, ale také varianční strukturu náhodných efektů (matice  $\mathbf{G}$  a  $\mathbf{R}$ ).

### Animal model na jednu vlastnost (single trait AM)

Jedná se o nejjednodušší smíšený model využívaných ve šlechtění zvířat. Určuje se plemenná hodnota pro každé zvíře, které má pouze jedno pozorování u jedné vlastnosti. Je zde tedy jen jeden pevný efekt a aditivně genetické efekty (nejsou zde žádné další náhodné efekty, jako jsou efekty maternální nebo dominance). Méně jednodušší modely jsou založeny na tomto základním modelu a nejsou tedy o moc více složitější pro pochopení.

V single trait AM pro odhad PH je pouze jeden náhodný efekt a předpokládáme, že matice  $\mathbf{R}$  je rovna  $\mathbf{I}\sigma_e^2$  a matice  $\mathbf{G}$  je  $\mathbf{I}\sigma_a^2$ . Po dosazení a vynásobením  $\sigma_e^2$  se normální rovnice zjednoduší na:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \quad \text{kde } \lambda = \frac{\sigma_e^2}{\sigma_a^2}$$

### Sire model – otcovský model

V tomto modelu jsou pouze efekty otců, které jsou odhadovány na základě záznamů jejich potomků (kteří jsou polovinou jejich plemenných hodnot!) a jsou výpočetně velmi snadné. Je potřeba méně rovnic než v animal modelu.

Koeficient  $\lambda$  je poměrem reziduální variance (ta zahrnuje  $\frac{3}{4}$  aditivně genetické variance) a variance otců ( $\frac{1}{4}$  aditivně genetické variance) a řešením jsou efekty otců, tj.  $\frac{1}{2}$  plemenných hodnot.

Odhadovaná PH je méně přesnější (méně potomků po otcích) a může být odhad vychýlený (není zde korekce pro rozdíly mezi matkami). Model předpokládá, že všichni potomci po otci jsou po různých matkách a všechny matky jsou vybrány ze stejné homogenní populace se stejným očekávaným průměrem. Matky však ve skutečnosti mohou pocházet z různých plemen, mohou být selektovány, kdy mladší matky mají pravděpodobně lepší genetické založení.

### Redukovaný animal model (RAM)

V tomto modelu jsou plemenné hodnoty odhadnuty pouze pro jedince, kteří mají pouze záznamy potomků. Tento model je rychlejší pro výpočet (jsou zde pouze rovnice pro zvířata, která jsou rodiči) a odhadované plemenné hodnoty pro všechna zvířata jsou jednoduše odvozena od OPH jejich rodičů plus jejich vlastní korigované fenotypy. Výsledky jsou stejné jako pro úplný animal model. Potřeba méně výpočetního času je narazena nutností programování navíc (což také spotřebuje čas výzkumníka).

### Model s opakovanými záznamy - odhad permanentních efektů prostředí

Tento model je používán, když u jednoho zvířete můžeme opakovaně změřit jeho užitek za jeho života. Fenotypová korelace mezi záznamy je rovna opakovatelnosti a předpokládá se, že genetická korelace mezi záznamy je rovna jedné (jestliže by byla genetická korelace menší než jedna, pak by se musel aplikovat víceznakový model).

Často je u zvířat sledována jedna vlastnost vícekrát. Jako příklad mohou sloužit:

- Hmotnost vlny ovcí v různých letech.
- Denní nádoj mléka v laktaci u dojného skotu.
- Velikost vrhu u prasnic.
- Velikost paroží u jelenů v různých sezónách.
- Výsledek závodu u koní z různých dostihů.

Kromě aditivně genetické hodnoty vlastnosti se zde uplatňuje efekt společného permanentního prostředí (PE), který je efektem negenetickým a společným pro všechna pozorování na stejném zvířeti.

Přístupem je zahrnout efekt permanentního prostředí pro každé zvíře, tj. kdy zvíře má druhý záznam, ne pouze svou PH, ale také část prostředových efektů, které se opakují. To může představovat efekty odchovu zvířete (dobré vývojové podmínky zaručí další dobrou užitkovost), nebo výskyt nemoci, která se projeví u konkrétního jedince, s permanentními efekty.

Model může být zapsán jako: 
$$\mathbf{y} = \mathbf{Xb} + (\mathbf{0} \ \mathbf{Z}) \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_r \end{pmatrix} + \mathbf{Zp} + \mathbf{e},$$

kde  $\mathbf{b}$  – vektor pevných efektů  
 $\mathbf{a}_0$  – zvířata bez záznamů  
 $\mathbf{a}_r$  – zvířata se záznamy  
 $\mathbf{p}$  – vektor PE efektů s délkou rovnou  $\mathbf{a}_r$   
 $\mathbf{e}$  – vektor reziduálních efektů

Maticе X a Z jsou designové matice asociované s pozorováním konkrétní úrovně pevných efektů a efektů aditivně genetických a PE. V modelu s opakovanými záznamy není Z rovno identické matici, takže,

$$\begin{aligned} \mathbf{a} | \mathbf{A}, \sigma_a^2 &\sim N(\mathbf{0}, \mathbf{A}\sigma_a^2) \\ \mathbf{p} | \mathbf{I}, \sigma_p^2 &\sim N(\mathbf{0}, \mathbf{I}\sigma_p^2) \\ \mathbf{e} &\sim N(\mathbf{0}, \mathbf{I}\sigma_e^2) \\ \mathbf{G} &= \begin{pmatrix} A\sigma_a^2 & 0 \\ 0 & \mathbf{I}\sigma_e^2 \end{pmatrix} \end{aligned}$$

Opakovatelnost je mírou pravděpodobnosti zvířete, že se na něm zopakuje vlastnost na stejné úrovni jako v předcházejícím záznamu (období, ...), a je definována jako poměr variancí:

$$r = \frac{\sigma_a^2 + \sigma_p^2}{\sigma_a^2 + \sigma_p^2 + \sigma_e^2}$$

### Smíšený model s opakovanými záznamy:

Model: 
$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Zp} + \mathbf{e}$$

kde  $\mathbf{b}$  je vektor pevných efektů s incidenční (designován) maticí X  
 $\mathbf{a}$  je vektor náhodných aditivně genetických efektů  
 $\mathbf{p}$  je vektor permanentních prostředových efektů  
 $\mathbf{Z}$  je incidenční matice, která se vztahuje k pozorováním u zvířat

Každé zvíře má aditivně genetický stejně jako permanentně prostředkový efekt, takže mají oba stejnou designovou matici.

Tři náhodné efekty mají následující distribuci:

$$\text{var} \begin{bmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & 0 & 0 \\ 0 & \mathbf{I}\sigma_c^2 & 0 \\ 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix} = \begin{bmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{R} \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & 0 \\ 0 & \mathbf{I}\sigma_c^2 \end{bmatrix}$$

Kde  $\sigma_a^2$  je přímá aditivně genetická variance a  $\sigma_c^2$  je variance způsobená permanentními prostředkovými efekty. Model ukazuje, že mezi zvířaty nejsou korelace v jejich aditivních a jejich permanentně prostředkovým efekty. Celková fenotypová variance je součtem tří komponent variancí.

Normální rovnice

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + k_1\mathbf{A}^{-1} & \mathbf{Z}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z} + k_2\mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

-  $k_1 = \sigma_e^2 / \sigma_a^2$  a  $k_2 = \sigma_e^2 / \sigma_c^2$

Př.: Tabulka níže obsahuje záznamy o velikosti vrhu šesti prasnic. Chceme odhadnout aditivně genetickou hodnotu (PH) těchto prasnic a jejich rodičů pro velikost vrhu pomocí AM BLUP.

prasnice	otec	matka	stádo	rok	pořadí vrhu	velikost vrhu	stádo-rok
3	?	?	1	1	2	10	1
4	1	?	1	1	1	6	1
4	1	?	1	2	2	7	2
5	2	3	2	1	1	8	3
5	2	3	2	2	2	13	4
6	?	3	2	1	1	12	3
6	?	3	1	2	2	10	2
7	2	6	2	2	1	10	4
8	1	3	1	2	1	5	2

Při OPH pro velikost vrhu předpokládáme model:

$$y_{ijklm} = \text{SR}_i + V_j + A_k + \text{PE}_l + e_{ijklm}$$

- SR - efekt stáda-roku
- V - efekt pořadí vrhu
- A - plemenná hodnota prasnice
- PE - permanentní prostředí prasnice
- e - temporální prostředí prasnice (reziduum)

Jedná se o animal model pro opakované pozorování, protože předpokládá další záznamy u stejné prasnice pro stejnou vlastnost. Tento model předpokládá, že velikost vrhu je vlastností prasnice než selat toho vrhu. Velikost vrhu je determinován geny prasnice.

Předpokládáme, že heritabilita je  $h^2 = 0,10$  a opakovatelnost  $r = 0,20$ .

Maticový zápis:

$$y = Xb + Z_1a + Z_2p + e$$

- $y$  – vektor pozorování
- $X$  – incidenční matice pro záznamy pevných efektů
- $b$  – neznámý vektor pevných efektů
- $Z_1$  – incidenční matice pro záznamy plemenných hodnot
- $a$  – neznámý vektor plemenných hodnot
- $Z_2$  – incidenční matice pro záznamy prostředových efektů
- $p$  – neznámý vektor permanentních prostředových efektů
- $e$  – vektor temporální prostředí prasnice (reziduum)

Prvky neznámých vektorů  $b$ ,  $u$  a  $p$  mohou být odhadnuty rovnicemi smíšeného modelu:

$$\begin{bmatrix} X'X & X'Z_1 & X'Z_2 \\ Z_1'X & Z_1'Z_1 + A^{-1}k_1 & Z_1'Z_2 \\ Z_2'X & Z_2'Z_1 & Z_2'Z_2 + Ik_2 \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ Z_1'y \\ Z_2'y \end{bmatrix}$$

$$\begin{array}{c} \mathbf{y} \\ \left[ \begin{array}{c} 10 \\ 6 \\ 7 \\ 8 \\ 13 \\ 12 \\ 10 \\ 10 \\ 5 \end{array} \right] \end{array} = \begin{array}{c} \mathbf{X} \\ \begin{array}{c} \overbrace{SR} \\ \underbrace{V} \\ \left[ \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{array} \right] \end{array} \end{array} \begin{array}{c} \mathbf{b} \\ \left[ \begin{array}{c} HY_1 \\ HY_2 \\ HY_3 \\ HY_4 \\ L_1 \\ L_2 \end{array} \right] \end{array} + \begin{array}{c} \mathbf{Z}_1 \\ \left[ \begin{array}{cccccccc} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \end{array} \begin{array}{c} \mathbf{PH} \\ \mathbf{a} \\ \left[ \begin{array}{c} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \\ A_6 \\ A_7 \\ A_8 \end{array} \right] \end{array} \\ \\ + \begin{array}{c} \mathbf{Z}_2 \\ \left[ \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right] \end{array} \begin{array}{c} \mathbf{p} \text{ PE} \\ \left[ \begin{array}{c} PE_3 \\ PE_4 \\ PE_5 \\ PE_6 \\ PE_7 \\ PE_8 \end{array} \right] \end{array} + [\mathbf{e}]$$

Sestavení pomocí tabulární metody matice aditivní příbuznosti  $A$  a invertovat  $A^{-1}$ :

$$A = \begin{bmatrix} 1 & 0 & 0 & 0,5 & 0 & 0 & 0 & 0,5 \\ 0 & 1 & 0 & 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0 & 1 & 0 & 0,5 & 0,5 & 0,25 & 0,5 \\ 0,5 & 0 & 0 & 1 & 0 & 0 & 0 & 0,25 \\ 0 & 0,5 & 0,5 & 0 & 1 & 0,25 & 0,375 & 0,25 \\ 0 & 0 & 0,5 & 0 & 0,25 & 1 & 0,5 & 0,25 \\ 0 & 0,5 & 0,25 & 0 & 0,375 & 0,5 & 1 & 0,125 \\ 0,5 & 0 & 0,5 & 0,25 & 0,25 & 0,25 & 0,125 & 1 \end{bmatrix}$$

Protože:

$$k_1 = \frac{\sigma_e^2}{\sigma_a^2} = \frac{1-r}{h^2} = 8 \quad k_2 = \frac{\sigma_e^2}{\sigma_{PE}^2} = \frac{1-r}{1-h^2} = 8$$

Matice koeficientů, matice levé strany rovnice je rovna:

SR				V		A								PE							
1	2	3	4	1	2	1	2	3	4	5	6	7	8	3	4	5	6	7	8		
2	0	0	0	1	1	0	0	1	1	0	0	0	0	1	1	0	0	0	0		
0	3	0	0	1	2	0	0	0	1	0	1	0	1	0	1	0	1	0	1		
0	0	2	0	2	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0		
0	0	0	2	1	1	0	0	0	0	1	0	1	0	0	0	1	0	1	0		
1	1	2	1	5	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1		
1	2	0	1	0	4	0	0	1	1	1	1	0	0	1	1	1	1	0	0		
0	0	0	0	0	0	14.67	0	4	-5.33	0	0	0	-8	0	0	0	0	0	0		
0	0	0	0	0	0	0	16	4	0	-8	4	-8	0	0	0	0	0	0	0		
1	0	0	0	0	1	4	4	19.67	0	-8	-5.33	0	-8	1	0	0	0	0	0		
1	1	0	0	1	1	-5.33	0	0	12.67	0	0	0	0	0	2	0	0	0	0		
0	0	1	1	1	1	0	-8	-8	0	18	0	0	0	0	0	2	0	0	0		
0	1	1	0	1	1	0	4	-5.33	0	0	16.67	-8	0	0	0	0	2	0	0		
0	0	0	1	1	0	0	-8	0	0	0	-8	17	0	0	0	0	0	1	0		
0	1	0	0	1	0	-8	0	-8	0	0	0	0	17	0	0	0	0	0	1		
1	0	0	0	0	1	0	0	1	0	0	0	0	0	9	0	0	0	0	0		
1	1	0	0	1	1	0	0	0	2	0	0	0	0	0	10	0	0	0	0		
0	0	1	1	1	1	0	0	0	0	2	0	0	0	0	0	10	0	0	0		
0	1	1	0	1	1	0	0	0	0	0	2	0	0	0	0	0	10	0	0		
0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	9	0		
0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	9		

Matice pravé strany rovnice je rovna:

Řešení soustavy rovnic:

$$\begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{bmatrix} = [\mathbf{LS}]^{-1} \cdot [\mathbf{PS}]$$

$$\begin{bmatrix} 16 \\ 22 \\ 20 \\ 23 \\ 41 \\ 40 \\ \hline 0 \\ 0 \\ 10 \\ 13 \\ 21 \\ 22 \\ 10 \\ 5 \\ \hline 10 \\ 13 \\ 21 \\ 22 \\ 10 \\ 5 \end{bmatrix}$$

Protože determinant LS je roven nule, nutno dodat podmínky řešitelnosti. Můžeme například smazat jeden sloupec a jeden řádek asociovaný s jednou úrovní pevného efektu – např. s druhým vrhem ( $V_2$ ). Je nutné také smazat 6. řádek ve vektoru PS.

Řešením soustavy upravených rovnic je:

$$\begin{bmatrix} \widehat{SR}_1 \\ \widehat{SR}_2 \\ \widehat{SR}_3 \\ \widehat{SR}_4 \\ \widehat{V}_1 \\ \hline \widehat{A}_1 \\ \widehat{A}_2 \\ \widehat{A}_3 \\ \widehat{A}_4 \\ \widehat{A}_5 \\ \widehat{A}_6 \\ \widehat{A}_7 \\ \widehat{A}_8 \\ \hline \widehat{PE}_3 \\ \widehat{PE}_4 \\ \widehat{PE}_5 \\ \widehat{PE}_6 \\ \widehat{PE}_7 \\ \widehat{PE}_8 \end{bmatrix} = \begin{bmatrix} 9.8263 \\ 8.3744 \\ 13.3015 \\ 13.3290 \\ -3.4505 \\ \hline -0.0693 \\ -0.0967 \\ 0.0776 \\ -0.1443 \\ -0.1077 \\ 0.2809 \\ 0.0937 \\ 0.0080 \\ \hline 0.0107 \\ -0.1462 \\ -0.1964 \\ 0.3213 \\ 0.0031 \\ 0.0076 \end{bmatrix}$$

Řešení vektoru zahrnuje odhady pro systematické efekty prostředí (stádo – rok a pořadí vrhu), efekty plemenných hodnot pro všechna zvířata a odhady permanentní efektů prostředí pro prasnice. Řešení pro efekty stádo-rok odhaduje efekt rozdílných managementů mezi dvěma stády nebo obdobím roků 1 a 2 ve stejném stádě. Tyto efekty mohou být použity k porovnání efektivnosti managementu ve stádech. Řešení z animal modelu jsou využitelná pro tyto cíle než pro porovnání průměrů stád, protože řešení pro

efekt stádo-rok z rovnice AM jsou upraveny pro genetické rozdíly mezi stády nebo roky. Řešení efektu stádo-rok jsou také upraveny pro další systematické prostředkové efekty. V našem příkladě jsou upravena pro skutečnost, že průměrná velikost vrhu pro stádo 1 (7,6 selat na vrh) zahrnuje 2 záznamy z 1. vrhů a tři z 2. vrhů, zatímco průměr stáda 2 (10,75) zahrnuje tři záznamy z 1. vrhů a pouze jeden 2. vrhu.

Z toho lze vyvodit, že management byl lepší ve 2. stádě než v prvním. V prvním roce management stáda 2 zvýšil velikost vrhu téměř o 3,5 selete (13,3 - 9,8) vzhledem ke stádu 1. Ve 2. roce se zlepšil management pro stádo 2 vůči stádu 1 o téměř 5 selat (13,33 - 8,37).

Pro stádo 1 byla odhadnuta úroveň managementu, který byl nepatrně horší v roce 2 než v roce 1 (rozdíl téměř o 1,5 selete: 9,83 - 8,37). Pro stádo 2 byl management téměř vyrovnán v obou letech (13,3 - 13,33).

Interpretace tohoto řešení je limitováno málo získanými informacemi, které byly zahrnuty do výpočtu. Proto i přesnost odhadu je nízká.

Odhad -3,45 pro první vrh  $V_1$  naznačuje, že průměrně první vrhy měly o 3,45 selete méně než druhé vrhy.

Odhady genetických rozdílů pro velikost vrhu mezi zvířaty jsou malá, jak naznačuje řešení pro  $A_i$ , které jsou očekávané na základě nízké heritability této vlastnosti a omezeného množství dat, které jsou k dispozici. Odhady permanentních efektů prostředí (negenetický vliv prasníc) jsou také malé v důsledku nízké hodnoty koeficientu opakovatelnosti této vlastnosti.

Přesnost odhadu  $A_i$  (PH) a  $PE_i$ :

Přesnost pro OPH	$r_{\hat{A}_i \hat{A}_i} = \sqrt{1 - d_{A_i} K_1}$	Přesnost pro PE	$r_{\hat{PE}_i \hat{PE}_i} = \sqrt{1 - d_{PE_i} K_2}$
0,1327		0,1867	
0,1143		0,3205	
0,2269		0,2843	
0,2955		0,3202	
0,2198		0,1929	
0,2854		0,1984	
0,1859			
0,1864			

Odhady PH mohou být použity k hodnocení zvířat pro genetické cíle. Např., když páříme prasnice se stejným otcem, pak dcery prasnice 6 budou mít očekávanou produkci o 0,43 selete ve vrhu více než dcery prasnice 4 (0,28 - (-0,144)).

Odhady PH spolu s odhady PE efekty mohou být použity k odhadům budoucí užítkovosti stejných prasníc, které by byly použity se záměrem brakování. Např. odhad v dalším vrhu prasnice 6 je odchylka od průměru stáda:  $\hat{A}_6 + \hat{PE}_6 = 0,28 + 0,32 = +0,60$  a prasnice 4:  $\hat{A}_4 + \hat{PE}_4 = -0,144 - 0,146 = -0,29$ .

A3	0,0883	A6	0,6022
A4	-0,2904	A7	0,0968
A5	-0,3042	A8	0,0155

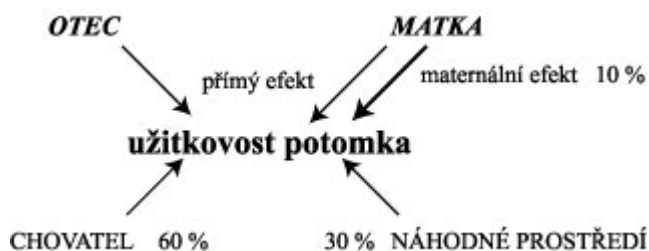
To znamená, že např. prasnice 4 a 6 budou-li mít svůj třetí vrh ve stejném stádě-roce, prasnice 6 bude mít velikost vrhu o 0,893 selete větší než prasnice 4 (0,6022 - (-0,2904)).



### Model s maternálními efekty

Vlastnosti jako jsou přežití selat (živě narozených) nebo časný růst u masných telat či jehňat jsou ovlivněny maternálním prostředím. Matka má vliv na užitkovost jejích potomků nad jejím přímým aditivně genetickým příspěvkem, tj. v důsledku maternálních efektů. Tyto efekty jsou pouze prostředím pro potomky, které však mohou být způsobeny genetickými i prostředovými složkami. Při selekci zvířat, a hlavně u matek, je důležité brát v úvahu maternální genetické efekty. Pro šlechtitele není důležité znát pouze plemennou hodnotu pro růst u masného skotu (přímý genetický efekt), ale také znát krávy s dobrými mateřskými schopnostmi (produkce mléka). Zahrnutím maternálního efektu do modelu nám dovoluje odhadnout maternální efekty a korigovat možné vychýlenosti v genetickém vyhodnocování rostoucích zvířat. Obvykle se předpokládá, že maternální efekt je genetický, i když jeho část může být efektem permanentního prostředí.

Řada vlastností je produkována jen matkami, např. přírůstky do odstavu. Tato vlastnost je ovlivněna polovinou genů, které tele získalo od své matky (**přímý genetický efekt**) a přírůstek je ovlivněn také kvalitou a množstvím mléka, chováním matky atd., které tele přijímá do odstavu - **nepřímý genetický efekt** neboli **materiální efekt** (mateřská způsobilost, která je determinována geneticky a prostředím).



Máme jednu naměřenou užitkovost, z které získáme dva odhady PH:

- PH vlastní užitkovosti
- PH maternálního efektu

Když  $y = A + E$ ,

a uvažujeme-li vliv materiální, lze rozčlenit  $E = E_D + A_M + E_M$ ,

kde  $E_D$  je přímý efekt prostředí (prostředí jiné než poskytuje matka);  $A_M$  je genetická hodnota pro mateřské schopnosti a  $E_M$  je prostředový efekt pro mateřskou schopnost.

Doplněním do celkové rovnice, kde zaměníme symbol  $A_D$  za  $A$  (přímý genetický efekt) získáme:

$$y = A_D + E_D + A_M + E_M$$

Protože kráva může mít za svůj život více telat, pak platí:

$$y = A_D + E_D + A_M + PE_M + TE_M$$

Efekty  $A_D$  a  $A_M$  získává potomek od rodiče a jsou důležité, když se selektuje na tuto vlastnost. Přímý genetický efekt zvířete ( $A_D$ ) je projeven v jeho vlastní užitkovosti. Nepřímý genetický efekt ( $A_M$ ) se projevuje v záznamech o užitkovosti jeho potomků.

Protože otcové hospodářských zvířat neodchovávají své potomky, otcovský  $A_M$  není nikdy projeven v záznamech užitkovosti svých potomků, ale dcery otce projeví geny pro mateřskou schopnost, kterou získali od otce ( $1/2 A_M$ ). Takže otcovský  $A_M$  je projeven v záznamech jeho dcer. Otec může být hodnocen pro maternální efekty, když jeho dcery projeví své mateřské schopnosti odchováním potomků.

Animal modely se mohou přizpůsobit pro hodnocení maternálních efektů, které dovolí oddělení přímých od maternálních genetických efektů a odhadnout PH pro přímý i maternální genetický efekt. Je-li přítomný, pak i pro odhad maternálních permanentních prostředových efektů.

### Smíšený model s maternálními efekty:

$$\text{Model: } \mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{m} + \mathbf{e}$$

kde  $\mathbf{b}$  je vektor pevných efektů s incidenční (designován) maticí  $\mathbf{X}$

$\mathbf{a}$  je vektor náhodných aditivně genetických efektů (přímý genetický efekt)

$\mathbf{m}$  je vektor maternálních genetických efektů

$\mathbf{Z}_1$  a  $\mathbf{Z}_2$  jsou incidenční matice, která se vztahují k náhodným efektům zvířete (aditivně genetický) a k matkám (maternálně genetický)

Tři náhodné efekty mají následující distribuci:

$$\text{var} \begin{bmatrix} \mathbf{a} \\ \mathbf{m} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & \mathbf{A}\sigma_{am} & 0 \\ \mathbf{A}\sigma_{am} & \mathbf{A}\sigma_m^2 & 0 \\ 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix} = \begin{bmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{R} \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & \mathbf{A}\sigma_{am} \\ \mathbf{A}\sigma_{am} & \mathbf{A}\sigma_m^2 \end{bmatrix} = \mathbf{G}_0 \otimes \mathbf{A}$$

Kde  $\mathbf{G}_0$  je matice 2 x 2:  $\begin{bmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix}$  a  $\otimes$  je přímý součin. Dále  $\sigma_a^2$  je přímá aditivně

genetická variance,  $\sigma_m^2$  je variance maternálně genetická,  $\sigma_{am}$  je kovariance mezi přímými a maternálně genetickými efekty a  $\sigma_e^2$  je reziduální variance. Model ukazuje, že oba náhodné efekty mají kovarianční strukturu závislou na genetické příbuznosti. Příbuzné matky mají příbuzné maternálně genetické efekty a vyskytují se korelace mezi přímým aditivně genetickým efektem matky a jejím maternálně genetickým efektem. Celková fenotypová variance je rovna  $\sigma_p^2 = \sigma_a^2 + \sigma_m^2 + \sigma_{am} + \sigma_e^2$ .

Normální rovnice smíšeného modelu:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_1 & \mathbf{X}'\mathbf{Z}_2 \\ \mathbf{Z}_1'\mathbf{X} & \mathbf{Z}_1'\mathbf{Z}_1 + k_{11}\mathbf{A}^{-1} & \mathbf{Z}_1'\mathbf{Z}_2 + k_{12}\mathbf{A}^{-1} \\ \mathbf{Z}_2'\mathbf{X} & \mathbf{Z}_2'\mathbf{Z}_1 + k_{21}\mathbf{A}^{-1} & \mathbf{Z}_2'\mathbf{Z}_2 + k_{22}\mathbf{A}^{-1} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{m}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}_1'\mathbf{y} \\ \mathbf{Z}_2'\mathbf{y} \end{bmatrix}$$

$$- \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} = \mathbf{G}_0^{-1} \cdot \sigma_e^2$$

Protože se používá pro výpočet aditivně genetická matice příbuznosti mezi všemi jedinci (potomci, matky i otcové), odhady budou získány pro aditivní efekty potomků (se záznamy užitkovosti) stejně jako pro matky i otce. Stejně jsou získány odhady maternálních efektů pro všechny jedince (nejen pro matky).

Pozn.: Počet odhadovaných genetických maternálních efektů je roven počtu jedinců v rodokmenu, zatímco počet efektů permanentního prostředí je roven počtu matek, které mají potomky s užitkovostí.

Př:

Tele	Otec	Matka	Pohlaví telete	Přírůstky (kg)
4	1	3	F	220
5	2	4	F	208
6	1	5	M	230
7	2	3	M	225

Model pro *přírůstek* do odstavu (kg) na *j*-té tele *i*-tého pohlaví od *k*-té matky:

$$y_{ijk} = p_i + AD_j + AM_k + PEM_k + TE_{Ink}$$

Maticový zápis:

$$y = Xb + Z_D A_D + Z_M A_M + Z_P P_M + e$$

$$\begin{bmatrix} 220 \\ 208 \\ 230 \\ 225 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} F \\ M \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_{D1} \\ A_{D2} \\ A_{D3} \\ A_{D4} \\ A_{D5} \\ A_{D6} \\ A_{D7} \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} A_{M1} \\ A_{M2} \\ A_{M3} \\ A_{M4} \\ A_{M5} \\ A_{M6} \\ A_{M7} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} PEM_3 \\ PEM_4 \\ PEM_5 \end{bmatrix} + [TE_k]$$

### Odhad PH pro více vlastností (Multi trait AM)

Matice  $R^{-1}$  musí zůstat v rovnici. Vektor  $y$  rozdělíme na více částí: kg mléka na 1. laktaci, % tuku a % bílkovin v mléce.

$$y = \begin{bmatrix} 5000 \\ 6000 \\ 7000 \\ 4,1 \\ 3,9 \\ 4,0 \\ 3,3 \\ 3,2 \\ 3,3 \end{bmatrix} \begin{array}{l} \text{- kg mléka} \\ \\ \\ \text{- \% tuku} \\ \\ \text{- \% bílkovin} \end{array}$$

- tomu se pak přizpůsobí ostatní matice (tzn. že jsou složeny ze tří částí dle vektoru  $y$  nutné používat matici  $R^{-1}$ )

Rozklad korelací na kovariance genetické a prostředkové.

$$\mathbf{R} = \begin{bmatrix} \text{cov}_E & 0 & 0 \\ 0 & \text{cov}_E & 0 \\ 0 & 0 & \text{cov}_E \end{bmatrix}$$

$\text{cov}_E$  – kovariance prostředková

- zde bloky 3 x 3 (3 vlastnosti)

Jestliže chybí nějaký naměřený údaj u jedné vlastnosti – díky korelacím lze předpovědět PH i v této vlastnosti. Nutné znát korelaci mezi vlastnostmi. Zde sledujeme 3 vlastnosti  $\Rightarrow$  pro každé zvíře nutno počítat 3 rovnice.

## SELEKČNÍ INDEXY

$$I = \sum b_i PH_i$$

$b_i$  – váhy indexu

$PH_i$  – plemenné hodnoty vlastnosti

$$I = b_1x + b_2y + \dots$$

$b$  – váhový koeficient

- zvíře je ve vlastnosti x nejlepší a ve vlastnosti y horší, proto tyto vlastnosti musíme zvážit koeficientem  $b$
- lze použít víceznakový AM, kde data  $x$  a  $y$  se berou jako čisté PH, ke kterým vkládáme dílčí PH – je to však složité
- nebo jednoduše odhadnout PH pro každou vlastnost zvlášť a pak každou zvážit koeficientem  $b$
- začleňují se také genetické a prostředkové kovarianční matice
- díky korelacím se dá předpovědět vlastnost

Předpoklad – data jsou očištěna od všech vlivů (tzn. stanovit OPH)

I stanovíme tak, že odhadneme celkovou hodnotu zvířete H:

$$H = a_1g_1 + a_2g_2 + \dots$$

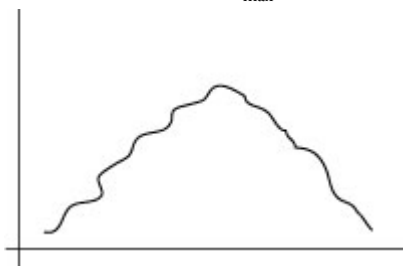
$a$  – cena za vlastnost 1 (ekonomická hodnota vlastnosti)

$g$  – genotyp vlastnosti 1

ALE  $g$  nelze zjistit, jen odhadnout. Šlechtí se pomocí  $g$ , takže se vyžaduje těsná korelace mezi  $H$  a  $I$  (maximální) -  $r_{\max}$ .

Použijeme průběh funkcí a hledáme extrém funkce korelačního koeficientu -  $r_{\max}$ :

$$r_{IH} = \frac{\text{COV}_{IH}}{\sigma_I \cdot \sigma_H}$$



$$\text{cov}_{IH} = a_1 b_1 \cdot \text{cov}_{g_1x} + a_1 b_2 \cdot \text{cov}_{g_1y} + a_2 b_1 \cdot \text{cov}_{g_2x} + a_2 b_2 \cdot \text{cov}_{g_2y} + \dots$$

Koeficienty indexu  $b$  ovlivňují velikost  $\text{cov}_{IH} \Rightarrow$  hledáme takové kombinace  $b_1, b_2$ , aby korelace  $r_{\max}$  byla co nejvyšší.

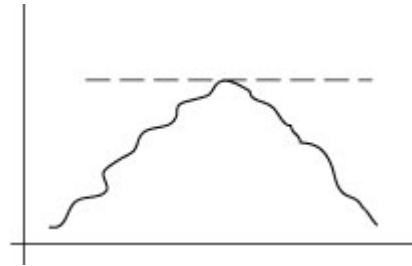
První parciální derivací získáme tečnu k funkci se směrnicí nulovou (vodorovnou) v maximum funkce.

- 1. parciální derivace = 0  
- hledáme nulovou směrnicí
- soustava rovnic, kterou řešíme
- získáme jednoduchý výpočet

$$P \cdot b = G \cdot a$$

$P$  – fenotypová kovarianční matice (1x1)

$G$  – genetická kovarianční matice (1x1)



**Př.** Selekční index o 1 vlastnosti a genotyp o 1 vlastnosti (mléčná užitkovost)

$$I = b \cdot X \quad H = a \cdot g \quad r_{\max}$$

$$[G] = 1 \times 1 = 200^2 \quad \text{- genotypový rozptyl}$$

$$P \cdot b = G \cdot a$$

$$[P] = 1 \times 1 = 400^2 \quad \text{- fenotypový rozptyl}$$

$$400^2 \cdot b = 200^2 \cdot 1$$

$$a = 1$$

$$b = 0,25$$

Má-li kráva odchylku 300 kg ( $PH = 300$ ), pak  $I = b \cdot PH = 0,25 \cdot 300 = 75$

$$\text{Spolehlivost odhadu PH: } r^2 = \frac{\text{var}_{PH}}{\text{var}_G} = \frac{\text{var}_I}{\text{var}_H} = 0,25$$

$$\text{var}_I = b' \cdot P \cdot b = 0,25' \cdot 400^2 \cdot 0,25 = 10000$$

$$\text{var}_H = a' \cdot G \cdot a = 1' \cdot 200^2 \cdot 1 = 40000$$

$r^2 = 0,25$  - protože selektujeme krávu podle 1 vlastnosti (nebyly brány v úvahu vrstevnice a předpokládalo se, že jich bylo mnoho, a tedy efektivní počet byl roven 1)

**Př.** Sledujeme 2 zvířata se 2 vlastnostmi:

vlastnost	1	2	a	$h^2$	$\sigma_G$	$r_G$	$r_P$
X	+ 500	+ 200	5	0,3	100	- 0,1	+ 0,1
Y	- 50	+ 10	12	0,4	50		

Sestavit genetickou a fenotypovou kovarianční matici<sup>1</sup>, náhodné koeficienty a odhadnout, které zvíře je lepší!

$G$  – na diagonále je genotypová variance a mimo diagonálu genotypová kovariance

$P$  – na diagonále je fenotypová variance a mimo diagonálu fenotypová kovariance

<sup>1</sup> Kovarianční matice – na diagonále jsou variance, mimo diagonálu jsou kovariance.

$$\sigma_{G_x}^2 = 10000 \quad \sigma_{P_x}^2 = \frac{\sigma_{G_x}^2}{h_x^2} = 33333,33$$

$$\sigma_{P_x} = 182,57$$

$$\sigma_{G_y}^2 = 2500 \quad \sigma_{P_y}^2 = \frac{\sigma_{G_y}^2}{h_y^2} = 6250$$

$$\sigma_{P_y} = 79,06$$

$$\text{cov}_{G_{xy}} = r_{G_{xy}} \cdot \sigma_{G_x} \cdot \sigma_{G_y}$$

$$\text{cov}_{G_{xy}} = -500,0$$

$$\text{cov}_{P_{xy}} = r_{P_{xy}} \cdot \sigma_{P_x} \cdot \sigma_{P_y}$$

$$\text{cov}_{P_{xy}} = 1443,4$$

$$h^2 = \frac{\sigma_G^2}{\sigma_P^2} \Rightarrow \sigma_P^2 = \frac{\sigma_G^2}{h^2}$$

$$r_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \cdot \sigma_y} \Rightarrow \text{cov}_{xy} = r_{xy} \cdot \sigma_x \cdot \sigma_y$$

$$G = \begin{bmatrix} \sigma_{G_x}^2 & \text{cov}_{G_{xy}} \\ \text{cov}_{G_{xy}} & \sigma_{G_y}^2 \end{bmatrix} = \begin{bmatrix} 10000 & -500 \\ -500 & 2500 \end{bmatrix} \quad P = \begin{bmatrix} \sigma_{P_x}^2 & \text{cov}_{P_{xy}} \\ \text{cov}_{P_{xy}} & \sigma_{P_y}^2 \end{bmatrix} = \begin{bmatrix} 33333,3 & 1443,4 \\ 1443,4 & 6250 \end{bmatrix}$$

$$a = \begin{bmatrix} 5 \\ 12 \end{bmatrix}$$

$$P \cdot b = G \cdot a$$

$$b = P^{-1} \cdot G \cdot a$$

$$b = \begin{bmatrix} 33333,3 & 1443,4 \\ 1443,4 & 6250 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 10000 & -500 \\ -500 & 2500 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 12 \end{bmatrix}$$

$$b = \begin{bmatrix} \frac{395135625}{1289060759} & -\frac{42084375}{1289060759} \\ -\frac{55536875}{368301074} & \frac{150098125}{368301074} \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 12 \end{bmatrix} = \begin{bmatrix} 1,14 \\ 4,14 \end{bmatrix}$$

$$I = b_1 \cdot X + b_2 \cdot Y = b' \cdot \begin{bmatrix} 500 & 200 \\ -50 & 10 \end{bmatrix} = \begin{bmatrix} 1,14 \\ 4,14 \end{bmatrix}' \cdot \begin{bmatrix} 500 & 200 \\ -50 & 10 \end{bmatrix} = \begin{bmatrix} 363,6 & 269,5 \end{bmatrix}$$

Pro chovatele má hodnotu 1. zvíře 364 Kč a 2. zvíře 270 Kč.

Materiály určené pro studenty specializace „Genetika a šlechtění hospodářských zvířat“ pro předmět **Genetika ve šlechtění zvířat** (letní semestr 2006).

Dr. Ing. Tomáš Urban; ÚMFGZ – pracoviště genetiky MZLU v Brně

<http://www.af.mendelu.cz/genetika/>

[urban@mendelu.cz](mailto:urban@mendelu.cz);

duben '06

© Urban 2006