Multiple Stage Selection

Jack Dekkers

Multi-trait breeding goal: $H = v_1g_1 + v_2g_2 + v_3g_3 + \ldots\ldots + v_ng_n = \mathbf{v'g}$

Information sources: $X_1, X_2, X_3, X_4, \ldots\ldots, X_m$

Multi-trait selection index: $I = b_1X_1 + b_2X_2 + b_3X_3 + \ldots\ldots + b_mX_m$

Optimal index weights: $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\,\mathbf{v}$

Selection on I maximizes response to selection in H, but requires all animals to be measured for all traits.

Multiple-stage selection:

Stage 1: select on $I_1 = b_1X_1 + b_2X_2 + \ldots + b_kX_k = \mathbf{b_1'X_1}$

Stage 2: select on $I_2 = b_1X_1 + b_2X_2 + \ldots + b_kX_k + b_{k+1}X_{k+1} + \ldots + b_mX_m = \mathbf{b_2'X}$

Only animals that are selected in stage 1 have to be evaluated for $X_{k+1}, \ldots, X_m$

- ➔ Cost savings
- ➔ Opportunities to increase population size for early stages

Optimal index weights:

Stage 1: I_1 : $\mathbf{b_1} = \mathbf{P_{11}}^{-1}\mathbf{G_1}\,\mathbf{v}$ $\mathbf{P_{11}} = \mathrm{Var}(\mathbf{X_1})$ $\mathbf{G_1} = \mathrm{Cov}(\mathbf{X_1}, \mathbf{g})$

Stage 2: I_2 : $\mathbf{b_2} = \mathbf{P}^{-1}\mathbf{G}\,\mathbf{v}$

$$\mathbf{P} = \begin{bmatrix} \mathbf{P_{11}} & \mathbf{P_{12}} \\ \mathbf{P_{21}} & \mathbf{P_{22}} \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \mathbf{G_1} \\ \mathbf{G_2} \end{bmatrix}$$

Optimal weights for index I_2 are not affected by selection on I_1 in stage 1, provided all data included in I_1 is also included in I_2 (Cunningham 1975 Theor. Appl. Genet. 46:55)

But accuracy and response to selection on I_2 **are** affected by selection on I_1:

Stage 1: accuracy of I_1: $r_1 = \sqrt{\dfrac{\mathbf{b_1'G_1v}}{\mathbf{v'Cv}}}$ Trait response vector: $\boldsymbol{S}_{g,1} = i_1 \dfrac{\mathbf{b_1'G_1}}{\sqrt{\mathbf{b_1'P_{11}b_1}}}$

Stage 2: accuracy of I_2: $r_2 = \sqrt{\dfrac{\mathbf{b_2'G^*v}}{\mathbf{v'C^*v}}}$ Trait response vector: $\boldsymbol{S}_{g,2} = i_2 \dfrac{\mathbf{b_2'G^*}}{\sqrt{\mathbf{b_2'P^*b_2}}}$

This assumes multi-variate normality of variables at stage 2 (despite stage 1 selection).

Total response vector across both stages: $\boldsymbol{S}_g = \boldsymbol{S}_{g,1} + \boldsymbol{S}_{g,2}$

Matrices $\mathbf{P}^*$, $\mathbf{G}^*$, and $\mathbf{C}^*$ are $\mathbf{P}$, $\mathbf{G}$, and $\mathbf{C}$ matrices adjusted for selection on I_1

Matrix equivalent of adjustment of (co-)variance for selection on variable w (used for Bulmer effect):

$$\sigma_{xy}^{*} = \sigma_{xy} - k\frac{\sigma_{wx}\sigma_{wy}}{\sigma_w^2} \qquad k = i(i-t) \qquad t = \text{truncation point}$$

Consider vectors **w**, **x**, **y** Select on index **b'w**

$$\text{Cov}(\mathbf{x},\mathbf{y})^{*} = \text{Cov}(\mathbf{x},\mathbf{y}) - k\,\frac{\text{Cov}(\mathbf{x},\mathbf{b'w})\,\text{Cov}(\mathbf{b'w},\mathbf{y})}{\text{Var}(\mathbf{b'w})}$$

$$= \text{Cov}(\mathbf{x},\mathbf{y}) - k\,\frac{\text{Cov}(\mathbf{x},\mathbf{w})\mathbf{bb'}\text{Cov}(\mathbf{w},\mathbf{y})}{\mathbf{b'}\text{Var}(\mathbf{w})\mathbf{b}}$$

With stage 1 selection on $\mathbf{b'w} = \mathbf{b_1'X_1}$ → Matrices to use in Stage 2:

$$\mathbf{P}^{*} = \text{Var}(\mathbf{X})^{*} = \text{Cov}(\mathbf{X},\mathbf{X})^{*} = \mathbf{P} - k\,\frac{\mathbf{Cov}(\underline{\mathbf{X}},\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\mathbf{Cov}(\underline{\mathbf{X}}_1,\underline{\mathbf{X}})}{\underline{\mathbf{b}}_1'\mathbf{Var}(\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1}$$

$$= \mathbf{P} - k\,\frac{\begin{bmatrix}\underline{\mathbf{P}}_{11}\\ \underline{\mathbf{P}}_{21}\end{bmatrix}\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\begin{bmatrix}\underline{\mathbf{P}}_{11} & \underline{\mathbf{P}}_{21}\end{bmatrix}}{\underline{\mathbf{b}}_1'\underline{\mathbf{P}}_{11}\underline{\mathbf{b}}_1}$$

$$\mathbf{G}^{*} = \text{Cov}(\mathbf{X},\mathbf{g})^{*} \qquad = \mathbf{G} - k\,\frac{\mathbf{Cov}(\underline{\mathbf{X}},\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\mathbf{Cov}(\underline{\mathbf{X}}_1,\underline{\mathbf{g}})}{\underline{\mathbf{b}}_1'\mathbf{Var}(\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1}$$

$$= \mathbf{G} - k\,\frac{\begin{bmatrix}\underline{\mathbf{P}}_{11}\\ \underline{\mathbf{P}}_{21}\end{bmatrix}\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\underline{\mathbf{G}}_1}{\underline{\mathbf{b}}_1'\underline{\mathbf{P}}_{11}\underline{\mathbf{b}}_1}$$

$$\mathbf{C}^{*} = \text{Var}(\mathbf{g})^{*} = \text{Cov}(\mathbf{g},\mathbf{g})^{*} = \mathbf{C} - k\,\frac{\mathbf{Cov}(\underline{\mathbf{g}},\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\mathbf{Cov}(\underline{\mathbf{X}}_1,\underline{\mathbf{g}})}{\underline{\mathbf{b}}_1'\mathbf{Var}(\underline{\mathbf{X}}_1)\underline{\mathbf{b}}_1}$$

$$= \mathbf{C} - k\,\frac{\underline{\mathbf{G}}_1'\underline{\mathbf{b}}_1\,\underline{\mathbf{b}}_1'\underline{\mathbf{G}}_1}{\underline{\mathbf{b}}_1'\underline{\mathbf{P}}_{11}\underline{\mathbf{b}}_1}$$

See 2-stage selection example.xls

Multi-stage selection with availability of multi-trait EBV:

EBV for all *m* traits available at every stage (but with different accuracies)

- select on complete index at every stage with weights = economic values

$$I = \mathbf{v_1 \hat{g}_1 + v_2 \hat{g}_2 + \ldots + v_n \hat{g}_n}$$

Optimization of proportions selected at each stage

Total proportion selected over *s* stages = P

Proportion selected at stage i = p_i

$$P = \prod_{i=1}^{s} \mathbf{p_i}$$

$\mathbf{a_i}$ = cost of traits measured at stage i

$$\text{Total cost} = \text{TC} = \mathbf{a_1} + \sum_{i=2}^{s} \mathbf{a_i} \prod_{j=1}^{i-1} \mathbf{p_j}$$

Maximize gain in breeding goal per unit of cost: Q = ΔH/TC

Or Maximize Profit from sale of breeding stock: Profit = N P k b ΔH – TC

N = total # animals in breeding program
P = percent selected
k = # times breed is multiplied before distribution
b = slope of supply-demand curve
= extra returns from sale of animal per extra unit genetic worth

Proportions selected and measured at each stage can then be optimized based on these objective functions and associated responses to selection.

References:

Cunningham (1975), Theor. Appl. Genet. 46:55
Ducrocq and Colleau (1989) Genet. Sel. Evol. 21:185
Xu and Muir (1991) Genetics 129:963
Xu and Muir (1992) Theor. Appl. Genet. 83:451
Xu, Martin, and Muir (1995) J. Anim. Sci. 73:699