A recipe for multiple trait deregression

I. Strandén and E.A. Mäntysaari

MTT Agrifood Research Finland

Biometrical Genetics

31600 Jokioinen

Ismo.Stranden@mtt.fi

Why this paper

- Original idea: Deregression into existing BLUP software
  - Easy to use
    - Get all benefits from existing software
  - Easy to program

- Deregression convergence
  - Convergence can be accelerated?
  - Many methods to choose

Base deregression equation system

Unknowns

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}+\mathbf{A}^{bb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ba}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{bg}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{ab}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{aa}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ag}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{gb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ga}\otimes\mathbf{G}_0^{-1} & \left(\mathbf{A}^{gg}+\mathbf{I}\right)\otimes\mathbf{G}_0^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{t}}_b \\ \hat{\mathbf{t}}_a \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_\mu \\ \mathbf{r}_b \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

b: Bulls with known EBVs ($\mathbf{a}_b$)

$$\hat{\mathbf{t}}_b = \mathbf{a}_b - \mathbf{X}\hat{\boldsymbol{\mu}}$$

a: Ancestors to bulls with known EBVs

g: Random genetic groups

Base deregression equation system

Unknowns

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}+\mathbf{A}^{bb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ba}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{bg}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{ab}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{aa}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ag}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{gb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ga}\otimes\mathbf{G}_0^{-1} & \left(\mathbf{A}^{gg}+\mathbf{I}\right)\otimes\mathbf{G}_0^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{t}}_b \\ \hat{\mathbf{t}}_a \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_\mu \\ \mathbf{r}_b \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

b: Bulls with known EBVs ($\mathbf{a}_b$)

$$\hat{\mathbf{t}}_b = \mathbf{a}_b - \mathbf{X}\hat{\boldsymbol{\mu}}$$

a: Ancestors to bulls with known EBVs

g: Random genetic groups

Relationship: $\mathbf{r}_b = \mathbf{R}^{-1}\mathbf{y}$ $\quad \mathbf{r}_\mu = \mathbf{X}'\mathbf{r}_b$

Deregressed EBVs

Block solving strategy: Step 1

Use current values

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}+\mathbf{A}^{bb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ba}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{bg}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{ab}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{aa}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ag}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{gb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ga}\otimes\mathbf{G}_0^{-1} & \left(\mathbf{A}^{gg}+\mathbf{I}\right)\otimes\mathbf{G}_0^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{t}}_b \\ \hat{\mathbf{t}}_a \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{\mu} \\ \mathbf{r}_b \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

b: Bulls with known EBVs ($\mathbf{a}_b$)

a: Ancestors to bulls with known EBVs

g: Random genetic groups

Solve for ancestors to bulls with known EBV and genetic groups

Needs solving a linear system of equations

Block solving strategy: Step 2

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}+\mathbf{A}^{bb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ba}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{bg}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{ab}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{aa}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ag}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{gb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ga}\otimes\mathbf{G}_0^{-1} & \left(\mathbf{A}^{gg}+\mathbf{I}\right)\otimes\mathbf{G}_0^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{t}}_b \\ \hat{\mathbf{t}}_a \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_\mu \\ \mathbf{r}_b \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

b: Bulls with known EBVs ($\mathbf{a}_b$)

a: Ancestors to bulls with known EBVs

g: Random genetic groups

Calculate right-hand side

Needs coefficient matrix times vector product

Block solving strategy: Step 3

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}+\mathbf{A}^{bb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ba}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{bg}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{ab}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{aa}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ag}\otimes\mathbf{G}_0^{-1} \\ \mathbf{0} & \mathbf{A}^{gb}\otimes\mathbf{G}_0^{-1} & \mathbf{A}^{ga}\otimes\mathbf{G}_0^{-1} & \left(\mathbf{A}^{gg}+\mathbf{I}\right)\otimes\mathbf{G}_0^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\mathbf{t}}_b \\ \hat{\mathbf{t}}_a \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_\mu \\ \mathbf{r}_b \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

b: Bulls with known EBVs ($\mathbf{a}_b$)

a: Ancestors to bulls with known EBVs

g: Random genetic groups

Calculate general mean: $\hat{\boldsymbol{\mu}}^{[k+1]} = \hat{\boldsymbol{\mu}}^{[k]} - \Delta^{[k]}$

k= iteration number

$$\Delta^{[k]} = \left(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{r}_\mu^{[k+1]} - \mathbf{X}'\mathbf{r}_b^{[k+1]}\right)$$

Block solving steps

1. Solve $\mathbf{t}_a$ and $\mathbf{g}$

   Needs solving a linear system of equations

   - Use existing BLUP solver: PCG iteration

2. Calculate new right-hand side

   - Matrix times vector product
     - Available: operation needed by PCG iteration

3. Update general mean

Iterate steps 1 to 3 until convergence

Accelerate solving of general mean

- Update in step 3:

$$\hat{\boldsymbol{\mu}}^{[k+1]} = \hat{\boldsymbol{\mu}}^{[k]} - \boldsymbol{\Delta}^{[k]}$$

$$\boldsymbol{\Delta}^{[k]} = \left(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{r}_{\mu}^{[k+1]} - \mathbf{X}'\mathbf{r}_{b}^{[k+1]}\right)$$

- Root finding method in update of general mean
  - Use $\boldsymbol{\Delta}^{[k]}$ as value for function at current general mean $\hat{\boldsymbol{\mu}}^{[k]}$

- Methods considered:
  - None
  - Bisection
  - Secant
  - Broyden

Acceleration by root finding methods

- Secant: $\boldsymbol{\mu}_i^{[k+1]} = \boldsymbol{\mu}_i^{[k]} - \frac{\boldsymbol{\mu}_i^{[k]} - \boldsymbol{\mu}_i^{[k-1]}}{\boldsymbol{\Delta}_i^{[k]} - \boldsymbol{\Delta}_i^{[k-1]}}$ for each trait i

- Broyden: $\boldsymbol{\mu}^{[k+1]} = \boldsymbol{\mu}^{[k]} - \mathbf{J}^{-1[k]}\boldsymbol{\Delta}^{[k]}$

$$\mathbf{J}^{-1[k]} = \mathbf{J}^{-1[k-1]} + \frac{\boldsymbol{\delta}^{[k]} - \mathbf{J}^{-1[k-1]}\boldsymbol{\Delta}^{[k]}}{\boldsymbol{\delta}^{[k]\prime}\mathbf{J}^{-1[k-1]}\boldsymbol{\Delta}^{[k]}}\boldsymbol{\delta}^{[k]\prime}\mathbf{J}^{-1[k-1]}$$

$$\boldsymbol{\delta}^{[k]} = \boldsymbol{\mu}^{[k]} - \boldsymbol{\mu}^{[k-1]}$$

- Extra computation due to acceleration is small

Block solving steps with acceleration

1. Solve $\mathbf{t}_a$ and $\mathbf{g}$

   Needs solving a linear system of equations

   - Use existing BLUP solver: PCG iteration

2. Calculate new right-hand side

   - Matrix times vector product
     - Available: operation needed by PCG iteration

3. Calculate function value at current general mean

$$\boldsymbol{\Delta}^{[k]} = \left(\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{r}_{\mu}^{[k+1]} - \mathbf{X}'\mathbf{r}_{b}^{[k+1]}\right)$$

4. Update general mean by acceleration method

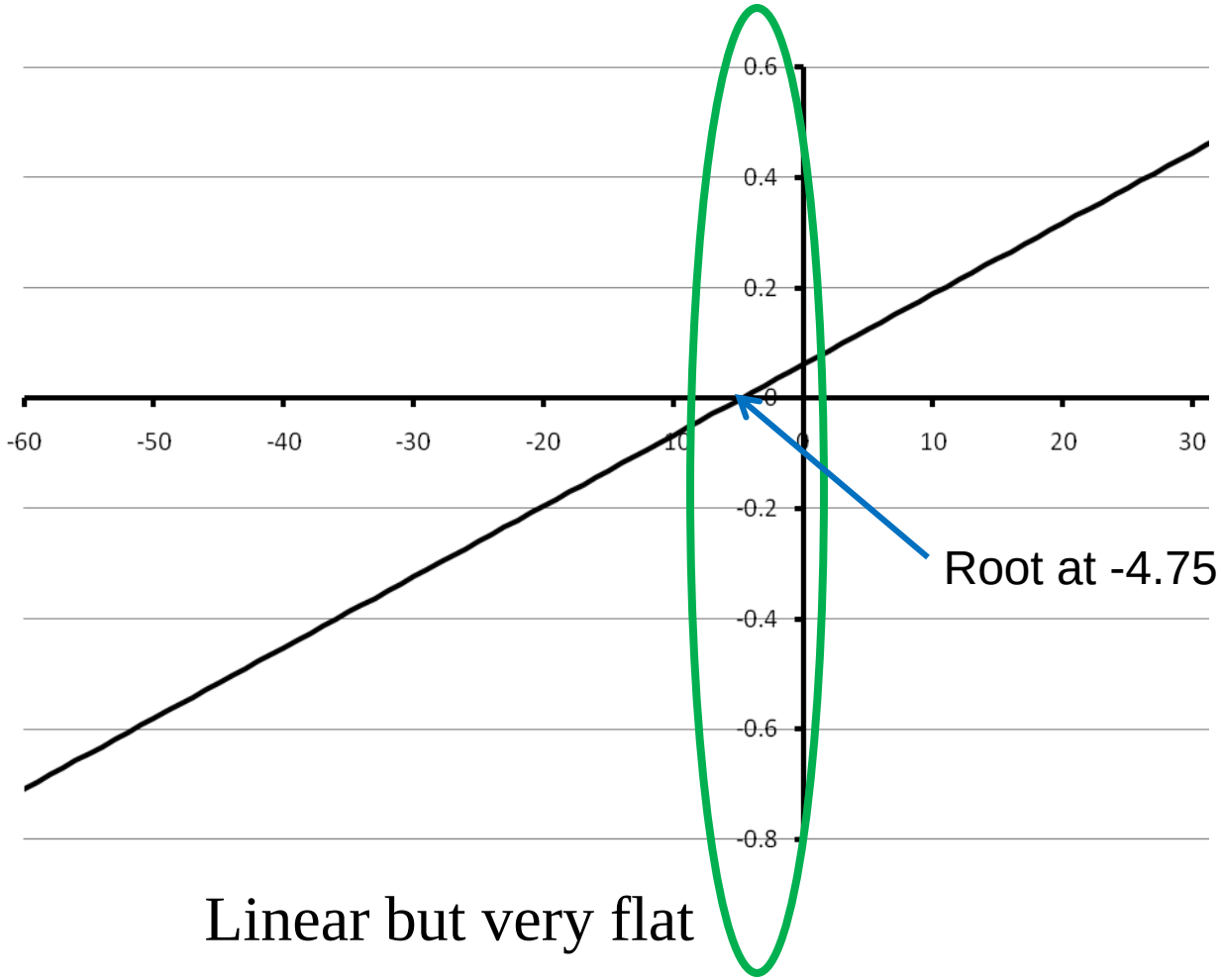
Iterate steps 1 to 4 until convergence

Data

- Two data sets from a paper by Schaeffer (2001)
- Country A
  - EBVs for 1st, 2nd, 3rd 305-d lactation protein yield
  - 4 analyses: 1, 1+2, 1+2+3 multiple trait,
    1+2+3 as single trait
- Country B
  - EBVs for protein yield and somatic cell score (SCS)
  - 4 analyses: protein, SCS, protein+SCS multiple trait,
    protein+SCS single trait

Results

Function value with different values of general mean, 305-d protein yield in country B



Iteration by different methods
Country B, protein

Iteration	None	Bisection	Secant	Broyden
0	-8.333	-13.5	-8.333	-8.333
1	-8.287	10.75	-8.287	-8.287
2	-8.242	-1.375	4126.96	-4.748
3	-8.197	-7.438	-4.748	-4.748
4	-8.153	-4.406	-4.748	-
No. iterations	713	16	4	3

Number of BLUP solver calls by analysis (total number of PCG iterations)

Country	Data	None	Bisection	Secant	Broyden
A	Lactation 1	138 (1656)	16 (191)	4 (48)	3 (36)
	MT_{1+2}	154 (1849)	250 (3093)	8 (97)	6 (73)
	MT_{1+2+3}	169 (2197)	269 (3497)	39 (506)	7 (91)
	All, ST[1]	138 (1794)	18 (233)	4 (51)	8 (104)
B	protein	713 (8556)	16 (192)	4 (49)	3 (36)
	SCS	149 (1506)	12 (128)	3 (47)	3 (34)
	$MT_{protein+SCS}$	748 (8976)	444 (5329)	6 (73)	6 (72)
	All ST[1]	713 (8556)	16 (193)	4 (49)	6 (72)

Number of BLUP solver calls by analysis (total number of PCG iterations)

Country	Data	None	Bisection	Secant	Broyden
A	Lactation 1	138 (1656)	16 (191)	4 (48)	3 (36)
B	protein	713 (8556)	16 (192)	4 (49)	3 (36)

Secant and Broyden methods are **very good**

Number of BLUP solver calls by analysis (total number of PCG iterations)

Country	Data	None	Bisection	Secant	Broyden
A	Lactation 1	138 (1656)	16 (191)	4 (48)	3 (36)
	MT_{1+2}	154 (1849)	250 (3093)	8 (97)	6 (73)
	MT_{1+2+3}	169 (2197)	269 (3497)	39 (506)	7 (91)
B	protein	713 (8556)	16 (192)	4 (49)	3 (36)
	SCS	149 (1506)	12 (128)	3 (47)	3 (34)
	$MT_{protein+SCS}$	748 (8976)	444 (5329)	6 (73)	6 (72)

Secant is **good**, and Broyden methods is **very good**

Number of BLUP solver calls by analysis (total number of PCG iterations)

Country	Data	None	Bisection	Secant	Broyden
A	Lactation 1	138 (1656)	16 (191)	4 (48)	3 (36)
	MT_{1+2}	154 (1849)	250 (3093)	8 (97)	6 (73)
	MT_{1+2+3}	169 (2197)	269 (3497)	39 (506)	7 (91)
	All, ST[1]	138 (1794)	18 (233)	4 (51)	8 (104)
B	protein	713 (8556)	16 (192)	4 (49)	3 (36)
	SCS	149 (1506)	12 (128)	3 (47)	3 (34)
	$MT_{protein+SCS}$	748 (8976)	444 (5329)	6 (73)	6 (72)
	All ST[1]	713 (8556)	16 (193)	4 (49)	6 (72)

Secant is **very good**, and Broyden methods is **good**

Number of PCG iterations by analysis

Country	Data	None	Bisection	Secant	Broyden
A	Lactation 1	138 (1656)	16 (191)	4 (48)	3 (36)
	MT_{1+2}	154 (1849)	250 (3093)	8 (97)	6 (73)
	MT_{1+2+3}	169 (2197)	269 (3497)	39 (506)	7 (91)
	All, ST[1]	138 (1794)	18 (233)	4 (51)	8 (104)
B	protein	713 (8556)	16 (192)	4 (49)	3 (36)
	SCS	149 (1506)	12 (128)	3 (47)	3 (34)
	$MT_{protein+SCS}$	748 (8976)	444 (5329)	6 (73)	6 (72)
	All ST[1]	713 (8556)	16 (193)	4 (49)	6 (72)

Results follow the same pattern as number of solver calls

Practical experiences (real data)

- Acceleration has worked very well
- Convergence affected by definition of genetic groups
  - The more groups the faster convergence
- Changes in group definition has only small effect on deregressed proofs (correlation)
  - Variance is affected
- Random genetic groups essential

Conclusions

- Deregression using existing BLUP software
  - Easy to implement
  - Gives all advantages of the existing software
    - User can start with deregression, and proceed to analyses with deregressed proofs easily:
      - same pedigree, same variance components etc.

- Acceleration methods work very well
  - No universally best among tested
  - Broyden's method was best when there were high genetic correlations between traits
  - Secant method was best when genetic correlations between traits were low