Talk on 2010-06-17

Model:

$$Y = X\gamma + H^A\beta^A + H^B\beta^B + G^{AB}\alpha + e$$

where

$$\beta^A \sim MN(0, \tau_A R_A)$$

$$\beta^B \sim MN(0, \tau_B R_B)]$$

$$\alpha \sim MN(0, \phi R_{AB})$$

$$e \sim MN(0, \sigma I)$$

Q1: Get the formula of REML for $\frac{\partial logL_{REML}}{\partial \phi}|_{\hat{\tau_A}, \hat{\tau_B}, \hat{\sigma}, \phi=0}$

A: Let $V = var(y) = \tau_A H^A R_A (H^A)^T + \tau_B H^B R_B (H^B)^T + \phi G^{AB} R^{AB} (G^{AB})^T + \sigma I$ according to REML (The detailed derivation see appendix):

$$L = -\frac{1}{2}(ln|V| + ln|X^T V^{-1} X| + Y^T P Y)$$

where $P = V^{-1} - V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}$ and

$$\frac{\partial logL_{REML}}{\partial \phi}|_{\hat{\tau_A}, \hat{\tau_B}, \hat{\sigma}, \phi=0} = Y^T P \frac{\partial V}{\partial \phi} Y - tr(\frac{\partial V}{\partial \phi} P)$$

It is easy to find out that $\frac{\partial V}{\partial \phi} = G^{AB}R^{AB}(G^{AB})^T$, so if τ_A, τ_B and σ are known, then $P_0 = P|_{\phi=0}$ is a constant. Therefore, the score test for ϕ is

$$T_\phi = \frac{1}{2} Y^T P_0 G^{AB} R^{AB} (G^{AB})^T Y - tr(G^{AB} R^{AB} (G^{AB})^T P)$$

Q2: One way to get $\hat{\tau_A}$, $\hat{\tau_B}$ and $\hat{\sigma}$ is to calculate the joint formulas:

$$\begin{cases} \frac{\partial logL_{REML}}{\partial \tau_A}|_{\phi=0} = 0 \\ \frac{\partial logL_{REML}}{\partial \tau_B}|_{\phi=0} = 0 \\ \frac{\partial logL_{REML}}{\partial \sigma}|_{\phi=0} = 0 \end{cases}$$

can we get the close forms of $\hat{\tau_A}$, $\hat{\tau_B}$ and $\hat{\sigma}$?

A:Define $S^A = H^A R_A (H^A)^T$ and $S^B = H^B R_B (H^B)^T$, we have

$$\begin{cases} \frac{\partial logL_{REML}}{\partial \tau_A}|_{\phi=0} = Y^T P_0 S^A - tr(S^A P_0) \\ \frac{\partial logL_{REML}}{\partial \tau_B}|_{\phi=0} = Y^T P_0 S^B - tr(S^B P_0) \\ \frac{\partial logL_{REML}}{\partial \sigma}|_{\phi=0} = Y^T P_0 - tr(P_0) \end{cases}$$

I am still working on how to decompose the $tr(.)$ to get a close form of $\hat{\tau_A}$, $\hat{\tau_B}$ and $\hat{\sigma}$.

Question: Since the close form seems hard to get, we actually performs a EM instead? Q3: The one Gene model is:

$$Y_i = X_i \gamma + H_i \beta + e_i$$

show that the score test from REML is the same with the score test derived from ML.

A:

$$\begin{aligned} L(\tau, \sigma) &= \prod f(Y_i|\tau, \sigma) \\ &= \prod \int f(Y_i, \hat{\gamma}|\tau, \sigma) d\beta \\ &= \prod \int f_y(Y_i|\beta, \tau, \sigma) f_\beta(\beta|\tau) d\beta \end{aligned}$$

APPENDIX

One of the annoying things about the ML is its estimation of variance does not take into account the loss in df for estimating the fixed effects. For example, if $y_i \sim N(\mu, \sigma^2)$, the estimator of σ^2 from ML is $\frac{\sum_1^n (y_i - \hat{\mu})^2}{n}$ which is biased, the unbiased estimator should be $\frac{\sum_1^n (y_i - \hat{\mu})^2}{n-1}$, that is, the MLE did not take the lost 1 degree of freedom which is used to estimate the μ into account. This is one of the motivations behind REML, which uses statistics that are unaffected by the fixed effects: $(\mathbf{I} - \mathbf{P_X})\mathbf{y}$. In other word, let $\hat{e}_i = y_i - \hat{u}$, then we will pick the $n-1$ independent $\hat{e}_i$ (the total n $\hat{e}_i$s are not mutually independent) rather than the total n independent y_i to estimate σ^2. By doing so, the REML estimator would be unbiased. We find a $n \times (n-p)$matrix $\mathbf{A}$ with condition $\mathbf{A^T A} = \mathbf{I_{n-p}}$ and $\mathbf{AA^T} = \mathbf{I} - \mathbf{P_X}$ and seek the distribution of $\mathbf{A^T y} \equiv \mathbf{w}$, which $\mathbf{w} \sim N_{n-p}(\mathbf{0}, \mathbf{A^T V A})$, which takes the form:

$$L_R(\theta) = (2\pi)^{-(n-p)/2} |\mathbf{A^T V A}|^{-1/2} \exp\{-\frac{1}{2} \mathbf{w^T} (\mathbf{A^T V A})^{-\mathbf{1}} \mathbf{w}\}.$$

To simplify the form above, we constructing the $n \times p$ matrix $\mathbf{G} = \mathbf{V^{-1}X(X^TV^{-1}X)^{-1}}$ and the $n \times n$ matrix $\mathbf{B} = [\mathbf{A}|\mathbf{G}]$. First the product,

$$\mathbf{B}^T\mathbf{B} = \begin{bmatrix} \mathbf{A}^T \\ \mathbf{G}^T \end{bmatrix} [\mathbf{A} \quad \mathbf{G}] = \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{G} \\ \mathbf{G}^T\mathbf{A} & \mathbf{G}^T\mathbf{G} \end{bmatrix}$$

then the determinant can be rewritten as

$$\begin{aligned} |\mathbf{B}|^2 = |\mathbf{B}'\mathbf{B}| &= |\mathbf{A}'\mathbf{A}| \times |\mathbf{G}'\mathbf{G} - \mathbf{G}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{G}| \\ &= |\mathbf{I}| \times |\mathbf{G}'\mathbf{G} - \mathbf{G}'(\mathbf{I} - \mathbf{P_X})\mathbf{G}| = |\mathbf{G}'P_{\mathbf{X}}\mathbf{G}| \\ &= |(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}| = |(\mathbf{X}'\mathbf{X})^{-1}|. \end{aligned}$$

which also says that $\mathbf{B}$ is nonsingular. Now following the similar steps, we find

$$\mathbf{B}'\mathbf{V}\mathbf{B} = \begin{bmatrix} \mathbf{A}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{V}[\mathbf{A} \quad \mathbf{G}] = \begin{bmatrix} \mathbf{A}^T\mathbf{V}\mathbf{A} & \mathbf{A}^T\mathbf{V}\mathbf{G} \\ \mathbf{G}^T\mathbf{V}\mathbf{A} & \mathbf{G}^T\mathbf{V}\mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{V}\mathbf{A} & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \end{bmatrix}$$

which yields only $|\mathbf{B}'\mathbf{V}\mathbf{B}| = |\mathbf{A}'\mathbf{V}\mathbf{A}|/|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|$. Some more manipulation puts all of the determinant result together:

$$\begin{aligned} |\mathbf{A}'\mathbf{V}\mathbf{A}| &= |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| \times |\mathbf{B}'\mathbf{V}\mathbf{B}| = |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| \times |\mathbf{B}|^2 \\ &= |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| \times |\mathbf{V}|/|\mathbf{X}'\mathbf{X}| \end{aligned}$$

For the quadratic form, the key is constructing the GLS estimator $\tilde{b} = \mathbf{G}'y = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}y$. Since $\mathbf{B}$ is nonsingular, we have

$$\begin{aligned} \mathbf{y}'\mathbf{V}^{-1}\mathbf{y} = \mathbf{y}'\mathbf{B}(\mathbf{B}'\mathbf{V}\mathbf{B})^{-1}\mathbf{B}'\mathbf{y} &= [w' \quad \tilde{b}'] \begin{bmatrix} \mathbf{A}'\mathbf{V}\mathbf{A} & 0 \\ 0 & (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \end{bmatrix} \begin{bmatrix} w \\ \tilde{b} \end{bmatrix} \\ &= w'(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}w + \tilde{b}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\tilde{b} \end{aligned}$$

so the quadratic form of interest can be rewritten as

$$\begin{aligned} w'(\mathbf{A}'\mathbf{V}\mathbf{A})^{-1}w &= y'\mathbf{V}^{-1}y - \tilde{b}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})\tilde{b} \\ &= y^TPy \end{aligned}$$

where $P = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}$