Aster Models for Life History Analysis

Charles J. Geyer
School of Statistics
University of Minnesota

Ruth G. Shaw and Helen H. Hangelbroek
Department of Ecology, Evolution, and Behavior
University of Minnesota

Stuart Wagenius
Institute for Plant Conservation Biology
Chicago Botanic Garden

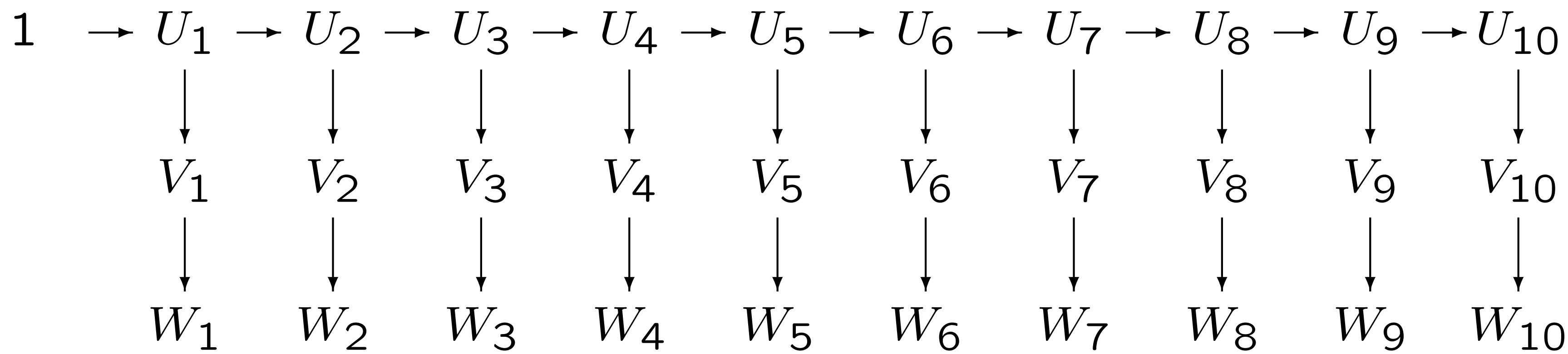
Julie R. Etterson
Department of Biology
University of Minnesota, Duluth

Geyer, C. J., Wagenius, S. and Shaw, R. G. (2007).
Aster models for life history analysis.
Biometrika, in press.
Technical Report No. 644 (all details of all computations)
http://www.stat.umn.edu/geyer/aster/

Shaw, R. G., Geyer, C. J., Wagenius, S., Hangelbroek, H. H., and Etterson, J. R. (submitted to *American Naturalist*).
Unifying life history analysis for inference of fitness and population growth.
Technical Report No. 658 (all details of all computations)
http://www.stat.umn.edu/geyer/aster/

Contributed package `aster` for R
http://www.r-project.org/

Graphical Model



U_i is conditionally Bernoulli given U_{i-1}.
U_1 is unconditionally Bernoulli.
V_i is conditionally Bernoulli given U_i.
W_i is conditionally zero-truncated Poisson given V_i.
(W_i is conditionally zero-inflated Poisson given U_i.)

U_i model survival, W_i model reproduction (number of offspring).

Aster Graphical Models

Simple graph structure. Each node has at most one parent.

Divide all nodes into root nodes F and non-root nodes J. Let $p(j)$ be parent of node j. Function p maps $J \to J \cup F$.

Graph indicates factorization of joint distribution as product of conditionals

$$\text{pr}(\mathbf{X}_J \mid \mathbf{X}_F) = \prod_{j \in J} \text{pr}(X_j \mid X_{p(j)})$$

Aster Model Families

Distribution of X_j given $X_{p(j)}$ is one-parameter exponential family with $X_{p(j)}$ as sample size.

Currently implemented families

- Bernoulli
- Poisson
- Negative binomial (known shape)
- Normal (known variance) — terminal nodes only
- k-truncated Poisson
- k-truncated negative binomial (known shape)

Log Likelihood, Conditional Parameterization

Log likelihood is simple

$$l(\boldsymbol{\theta}) = \sum_{j \in J} x_j \theta_j - x_{p(j)} c_j(\theta_j)$$

where c_j is cumulant function for one-parameter exponential family for j-th node. Canonical parameter of one-parameter exponential family is θ_j.

Extension to multi-parameter exponential families possible. Given in Geyer, Wagenius, and Shaw (2007), but not yet implemented.

Virtues of Aster Models

Joint analysis of all variables much better than separate analysis of GLM bits.

Separate analyses have "missing data" where joint aster analysis has only "structural zeros" (e. g., dead individual remains dead and cannot reproduce) which is built into the model structure — if $X_{p(j)}$ is zero, then distribution of X_j is concentrated at zero.

Structural zeros are handled automatically. Maximum likelihood just works.

Missing data present intractable problems.

Log Likelihood, Unconditional Parameterization

$$l(\boldsymbol{\theta}) = \sum_{j \in J} x_j \theta_j - x_{p(j)} c_j(\theta_j)$$

Collect terms with same x_j

$$\sum_{j \in J} x_j \left[\theta_j - \sum_{\substack{i \in J \\ p(i) = j}} c_i(\theta_i) \right] - \sum_{\substack{j \in J \\ p(j) \in F}} x_{p(j)} c_j(\theta_j)$$

Log Likelihood, Unconditional Parameterization (Cont.)

Recognize *unconditional exponential family* with new *canonical statistic vector* same as old (components x_j), new *canonical parameter vector* with components

$$\varphi_j = \theta_j - \sum_{\substack{i \in J \\ p(i) = j}} c_i(\theta_i) \qquad j \in J, \tag{$*$}$$

and new *cumulant function*

$$c(\boldsymbol{\varphi}) = \sum_{\substack{j \in J \\ p(j) \in F}} x_{p(j)} c_j(\theta_j)$$

Equations ($*$) define invertible change of parameter.

Canonical Affine Models and Dimension Reduction

Affine transformation

$$\boldsymbol{\varphi} = \mathbf{a} + \mathbf{M}\boldsymbol{\beta},$$

where $\mathbf{a}$ is known vector, $\mathbf{M}$ is known matrix, and $\boldsymbol{\beta}$ is unknown parameter vector, gives new exponential family with log likelihood

$$l(\boldsymbol{\beta}) = \mathbf{x}^T \mathbf{M}\boldsymbol{\beta} - c(\mathbf{a} + \mathbf{M}\boldsymbol{\beta})$$

new *canonical statistic vector* $\mathbf{M}^T\mathbf{x}$ and new *canonical parameter vector* $\boldsymbol{\beta}$.

Canonical statistic ($\mathbf{x}$ for full family, $\mathbf{M}^T\mathbf{x}$ for new subfamily) is *minimal sufficient*.

Virtues of Unconditional Parameterization

If sufficient statistic vector $\mathbf{y} = \mathbf{M}^T \mathbf{x}$ is scientifically interpretable this leads to direct interpretation of regression coefficients

$$-\frac{\partial^2 l(\boldsymbol{\beta})}{\partial \beta_k^2} = \frac{\partial E_{\boldsymbol{\beta}}\{Y_k\}}{\partial \beta_k} > 0.$$

Increasing one beta increases the unconditional expectation of corresponding canonical statistic, other betas being held fixed.

Intuitions derived from LM and GLM still valid, but only for unconditional parameterization.

Aster Model Printout for Example

```
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.176176   0.079214  -2.224 0.026145 *
vtypev      -1.422849   0.134546 -10.575  < 2e-16 ***
vtypew       1.349977   0.080444  16.782  < 2e-16 ***
uyear        0.041360   0.012163   3.400 0.000673 ***
z1           0.037439   0.007809   4.794 1.63e-06 ***
z2           0.021237   0.007101   2.990 0.002786 **
I(z1^2)     -0.032559   0.005350  -6.085 1.16e-09 ***
I(z2^2)     -0.026294   0.004815  -5.460 4.75e-08 ***
I(z1 * z2)   0.036837   0.008088   4.554 5.26e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

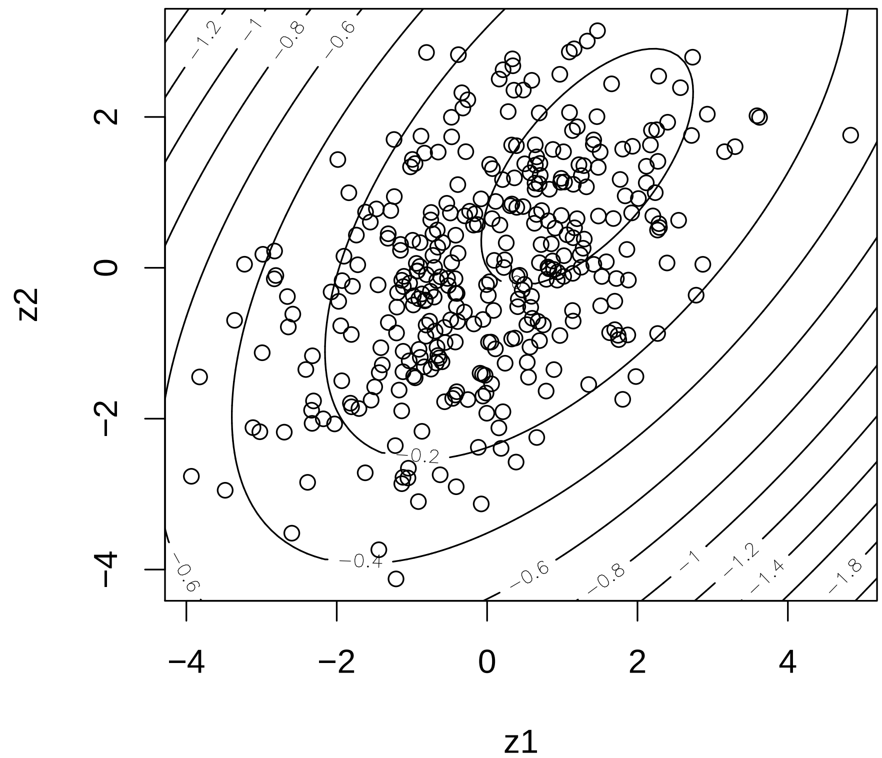
More About Model for Example

Let u_j^* be one if the j-th variable is a "U" and zero otherwise. Similarly for v_j^* and w_j^*.
Let t_j^* be the "time" (1, ..., 10) for the j-th variable.

$$
\begin{aligned}
\varphi_j = \beta_1 + u_j^*(\beta_2 + \beta_4 t_j^*) + \beta_3 v_j^* \\
+ w_j^*(\beta_5 z_1 + \beta_6 z_2 + \beta_7 z_1^2 + \beta_8 z_2^2 + \beta_9 z_1 z_2)
\end{aligned}
$$

z_1 and z_2 are covariates.

Monotone Transformation of Fitness Landscape



Mean Value Parameters

As in any exponential family

$$E_{\boldsymbol{\varphi}}(X) = \nabla c(\boldsymbol{\varphi})$$
$$\text{var}_{\boldsymbol{\varphi}}(X) = \nabla^2 c(\boldsymbol{\varphi})$$

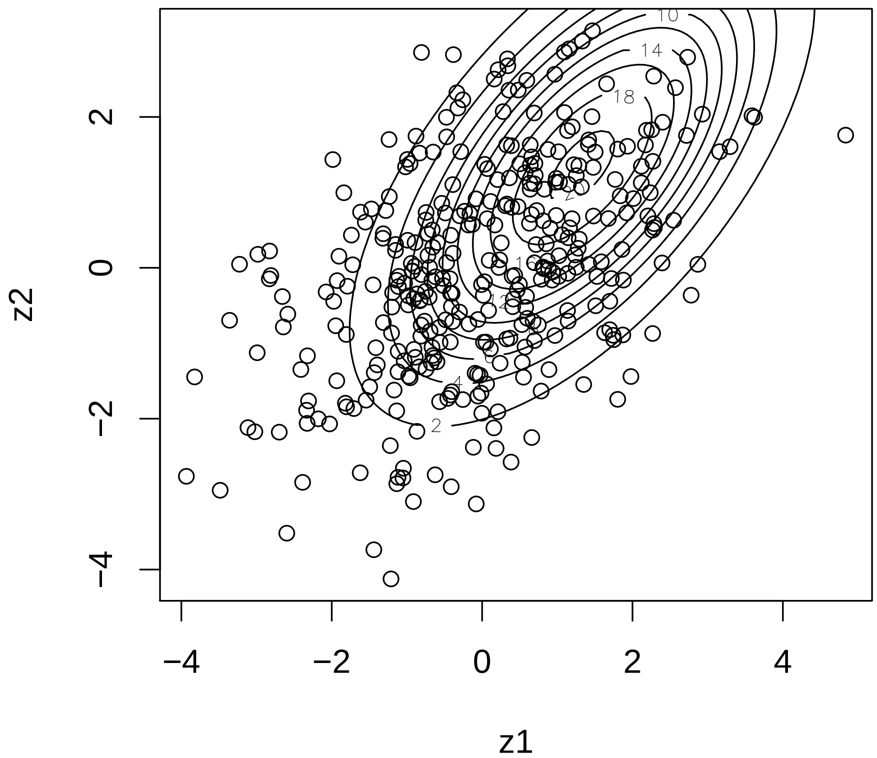
Since variance matrices are positive definite the mapping

$$\boldsymbol{\tau} : \boldsymbol{\varphi} \mapsto E_{\boldsymbol{\varphi}}(X)$$

is an invertible invertible change of parameter.

$\boldsymbol{\tau}$ is the *mean value parameter* of the full model.
$M^T \boldsymbol{\tau}$ is the mean value parameter of the affine submodel.

Fitness Landscape



Other Applications

Comparison of fitness among different groups

Estimation of population growth rate

Anything that fits aster model structure — very general

The Name of the Game

Photo from http://en.wikipedia.org/wiki/Sunflower

Aster models are named after the family *Asteraceae*, (type genus *Aster*, common name aster), which is huge (20,000 species).

They also come with a neat motto: *per aspera cum astris*, a take-off on the motto of the sunflower state.

Credit Where Credit is Due

co-authors

Janis Antonovics (R. Shaw's thesis advisor)

NSF (DMS-0083468, DEB-0545072, DEB-0544970) and EPA

Echinacea project http://echinacea.umn.edu, David Heiser, and Stacey Halpern

University of Minnesota Center for Community Genetics

The R project (www.r-project.org), especially Robert Gentleman
Sweave (Friedrich Leisch)