Pairwise Likelihood All Params

Sam Morris

28 May 2015

Likelihood for the data

The bivariate likelihood for the data is given as

$$f(Y_1, Y_2) = \begin{cases} 1 - \exp\left\{-\frac{1}{z_1}\right\} - \exp\left\{-\frac{1}{z_2}\right\} + \exp\left\{-\vartheta(\mathbf{s}_1, \mathbf{s}_2)\right\} & Y_1 = 1, Y_2 = 1 \\ \exp\left\{-\frac{1}{z_2}\right\} - \exp\left\{-\vartheta(\mathbf{s}_1, \mathbf{s}_2)\right\} & Y_1 = 1, Y_2 = 0 \\ \exp\left\{-\frac{1}{z_1}\right\} - \exp\left\{-\vartheta(\mathbf{s}_1, \mathbf{s}_2)\right\} & Y_1 = 0, Y_2 = 1 \\ \exp\left\{-\vartheta(\mathbf{s}_1, \mathbf{s}_2)\right\} & Y_1 = 0, Y_2 = 0 \end{cases} \tag{1}$$

where $z_i = \left(1 - \xi \mathbf{X}_i^T \beta\right)^{1/\xi}$, and $\vartheta(\mathbf{s}_1, \mathbf{s}_2) = \sum_{l=1}^{L} \left[\left(\frac{w_l(\mathbf{s}_1)}{z_1}\right)^{1/\alpha} + \left(\frac{w_l(\mathbf{s}_2)}{z_2}\right)^{1/\alpha}\right]^{\alpha}$

Pairwise composite likelihood with fixed ρ

Fixing all terms with MLE estimates. We found that things were not quite as good as probit and logit when the knots are spaced too far apart. For now, am just going to focus on the knot spacing that is a little bit closer than the truth. I am generating two datasets now to compare across multiple datasets.

Data settings

Right now, we're fitting $n = 2000$ observations with one replication. In the future, it would be nice to allow for multiple replications.

$\alpha = 0.25, \pi = 0.05, \rho = 0.1$

Pairwise likelihood estimates

Set 1

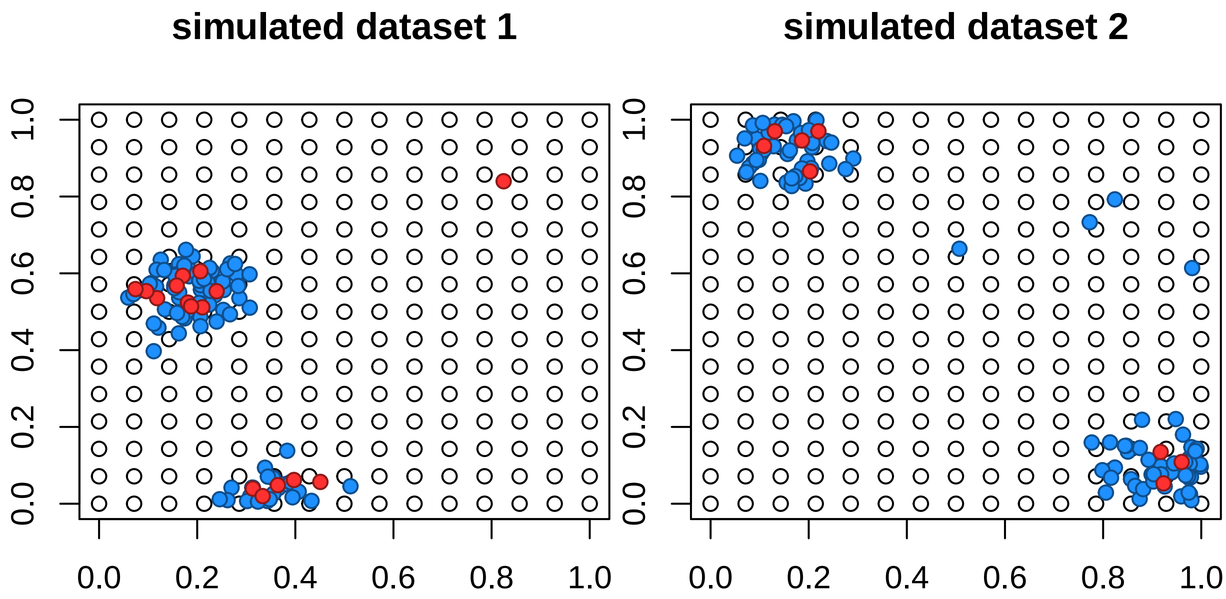
The true values are $\beta_0 = -5.5242711$, $\xi = 0.25$, and $\alpha = 0.25$. Our estimates are $\widehat{\alpha} = 0.27$, $\widehat{\xi} = 0.15$, and $\widehat{\beta} = -3.76$ when $\rho = 0.07$

Set 2

The true values are $\beta_0 = -3.399708$, $\xi = 0.25$, and $\alpha = 0.25$. The estimates are $\widehat{\alpha} = 0.38$, $\widehat{\xi} = 0.17$, and $\widehat{\beta} = -3.7$ when $\rho = 0.07$

Setup MCMC

We'll start with 1700 training sites, and 300 testing sites. In the following plot, the testing sites are given in red, the training sites are given in blue, and the knots are given as empty circles. In the simulated data, the knots are placed on a 12×12 grid.



Brier Scores

The brier scores are Logit 1: 0.0153 Logit 2: 0.005 Probit 1: 0.0149 Probit 2: 0.0047 GEV 1: 0.0177 GEV 2: 0.004

Generating another dataset

$\alpha = 0.25, \pi = 0.01, \rho = 0.1$

Pairwise likelihood estimates

Set 1

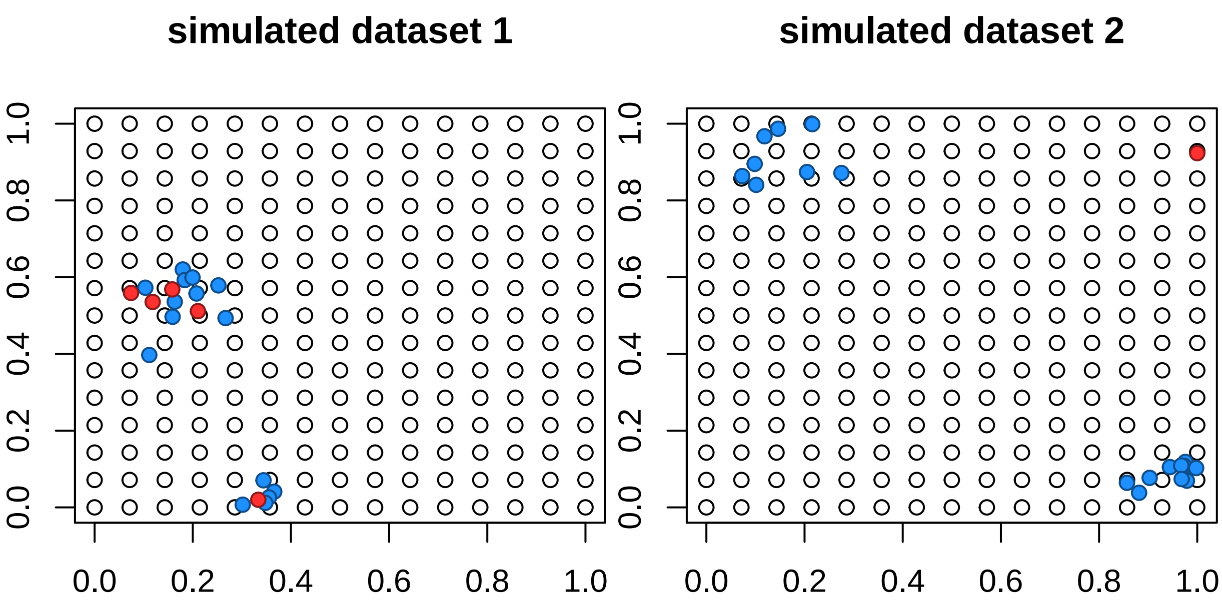
The true values are $\beta_0 = -6.9852927$, $\xi = 0.25$, and $\alpha = 0.25$. Our estimates are $\widehat{\alpha} = 0.83$, $\widehat{\xi} = -0.07$, and $\widehat{\beta} = -4.05$ when $\rho = 0.07$

Set 2

The true values are $\beta_0 = -4.6124884$, $\xi = 0.25$, and $\alpha = 0.25$. The estimates are $\widehat{\alpha} = 0.72$, $\widehat{\xi} = -0.04$, and $\widehat{\beta} = -4.12$ when $\rho = 0.07$

Setup MCMC

We'll start with 1700 training sites, and 300 testing sites. In the following plot, the testing sites are given in red, the training sites are given in blue, and the knots are given as empty circles. In the simulated data, the knots are placed on a 12×12 grid.



Brier Scores

The brier scores are Logit 1: 0.0165 Logit 2: 0.0024 Probit 1: 0.0129 Probit 2: 0.0021 GEV 1: 0.0139 GEV 2: 0.0023

Generating another dataset

$\alpha = 0.75, \pi = 0.05, \rho = 0.1$

Pairwise likelihood estimates

Set 1

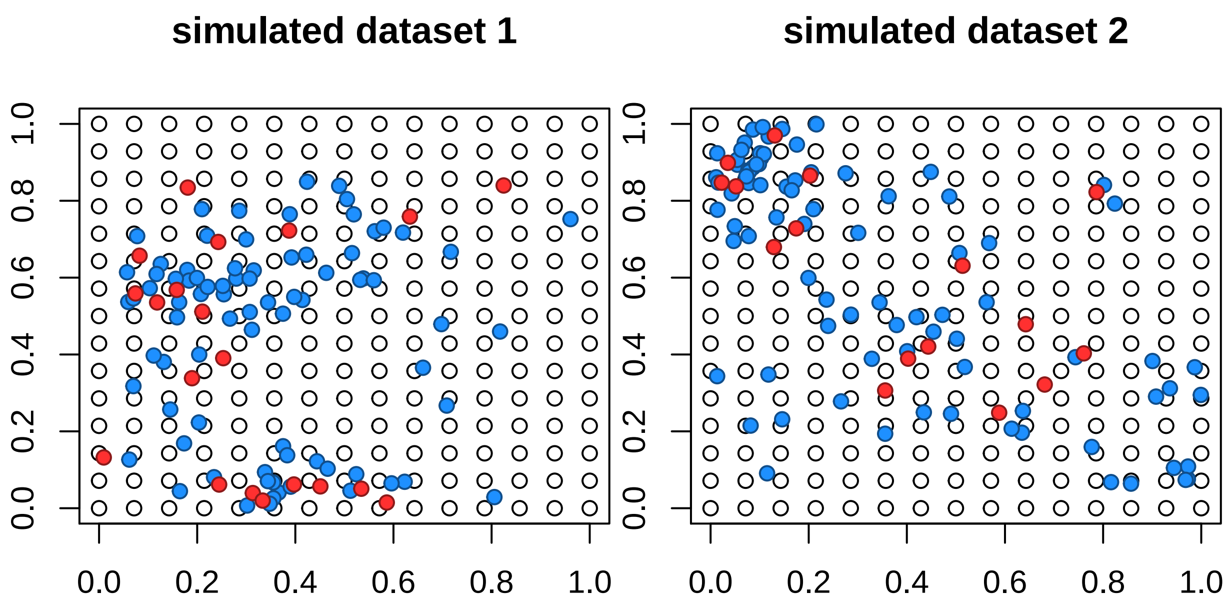
The true values are $\beta_0 = -4.4157785$, $\xi = 0.25$, and $\alpha = 0.75$. Our estimates are $\widehat{\alpha} = 0.89$, $\widehat{\xi} = 0.19$, and $\widehat{\beta} = -4.08$ when $\rho = 0.07$

Set 2

The true values are $\beta_0 = -3.632659$, $\xi = 0.25$, and $\alpha = 0.75$. The estimates are $\widehat{\alpha} = 0.84$, $\widehat{\xi} = 0.19$, and $\widehat{\beta} = -4.01$ when $\rho = 0.07$

Setup MCMC

We'll start with 1700 training sites, and 300 testing sites. In the following plot, the testing sites are given in red, the training sites are given in blue, and the knots are given as empty circles. In the simulated data, the knots are placed on a 12×12 grid.



Brier Scores

The brier scores are Logit 1: 0.0602 Logit 2: 0.0445 Probit 1: 0.0569 Probit 2: 0.0449 GEV 1: 0.0571 GEV 2: 0.0453

Generating another dataset

$\alpha = 0.75, \pi = 0.01, \rho = 0.1$

Pairwise likelihood estimates

Set 1

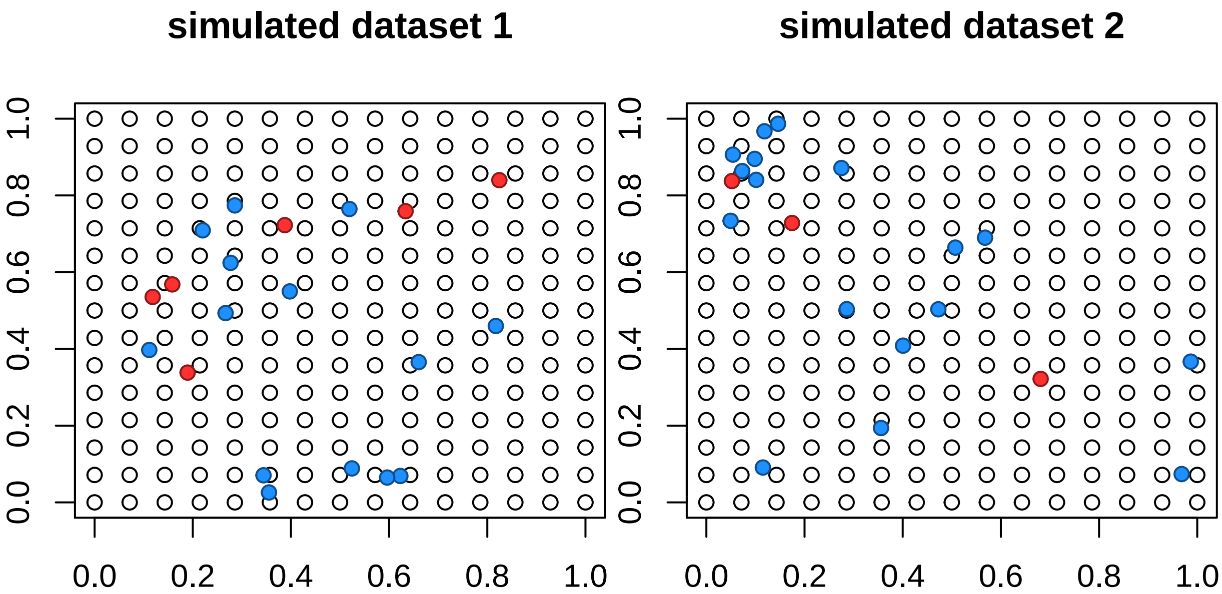
The true values are $\beta_0 = -7.1169039$, $\xi = 0.25$, and $\alpha = 0.75$. Our estimates are $\widehat{\alpha} = 1$, $\widehat{\xi} = -0.15$, and $\widehat{\beta} = -3.46$ when $\rho = 0.07$

Set 2

The true values are $\beta_0 = -5.920303$, $\xi = 0.25$, and $\alpha = 0.75$. The estimates are $\widehat{\alpha} = 0.96$, $\widehat{\xi} = -0.01$, and $\widehat{\beta} = -4.46$ when $\rho = 0.07$

Setup MCMC

We'll start with 1700 training sites, and 300 testing sites. In the following plot, the testing sites are given in red, the training sites are given in blue, and the knots are given as empty circles. In the simulated data, the knots are placed on a 12×12 grid.



Brier Scores

The brier scores are Logit 1: 0.0197 Logit 2: 0.0099 Probit 1: 0.0198 Probit 2: 0.0097 GEV 1: 0.3174 GEV 2: 0.0097