# Dirichlet process mixture model on Hopkinson-bar experiments

Mickey Warner

7 December 2015

AMS 241

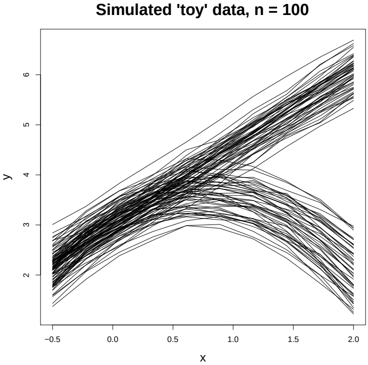
## Data: simulated

Generated $n = 100$ curves, each of length $k_i$,

$$\mathbf{y}_i = \beta_{i0}\mathbf{1}_{k_i} + \beta_{i1}\mathbf{x}_i + \beta_{i2}\mathbf{x}_i^2 + \boldsymbol{\epsilon}_i$$

where

- $\boldsymbol{\epsilon}_i \overset{iid}{\sim} N(0, 0.05^2\mathbf{I}_{k_i})$
- $\beta_{i0} \overset{iid}{\sim} N(3, 0.3^2)$
- $\beta_{i1} \overset{iid}{\sim} N(1.5, 0.1^2)$
- $\beta_{i2} \overset{iid}{\sim} 0.5\delta_0(\cdot) + 0.5N(-1, 0.1^2)$



## Data: simulated (continued)

We use $\mathbf{x}_1 = \cdots = \mathbf{x}_n = \mathbf{x}$ that is a vector of $k_i = 10$ equally spaced values from $-0.5$ to $2$

Thus we have 100 curves with random intercepts, slopes, and quadratic terms, about half of which are lines and the others are parabolas

# Parametric hierarchical model

The curves $\mathbf{y}_i$ are modeled as a multivariate normal with a quadratic polynomial as the mean:

$$\begin{aligned} f(\mathbf{y}_i|\mathbf{x}_i, \boldsymbol{\beta}_i, \tau^2) &= N_{k_i}(\mathbf{y}_i|p(\mathbf{x}_i, \boldsymbol{\beta}_i), \tau^2\mathbf{I}_{k_i}), & i = 1, \ldots, n \\ \boldsymbol{\beta}_i &\overset{iid}{\sim} N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), & i = 1, \ldots, n \\ \tau^2 &\sim IG(a_\tau, b_\tau) \\ \boldsymbol{\mu} &\sim N(\mathbf{m}, \mathbf{S}) \\ \boldsymbol{\Sigma} &\sim IW(\mathbf{V}, d) \end{aligned}$$

where $\boldsymbol{\beta}_i = (\beta_{i0}, \beta_{i1}, \beta_{i2})^\top$, $p(\mathbf{x}_i, \boldsymbol{\beta}_i) = \beta_{i0}\mathbf{1}_{k_i} + \beta_{i1}\mathbf{x}_i + \beta_{i2}\mathbf{x}_i^2$, and $a_\tau, b_\tau, \mathbf{m}, \mathbf{S}, \mathbf{V}, d$ are specified.

# Dirichlet process mixture (DPM) model

We mix over the parameters in the mean function to obtain

$$f(\mathbf{y}_i | G, \mathbf{x}_i, \tau^2) = \int N_{k_i}(\mathbf{y}_i | p(\mathbf{x}_i, \boldsymbol{\beta}), \tau^2 \mathbf{I}_{k_i}) dG(\boldsymbol{\beta})$$

$$G | \alpha, G_0 \sim DP(\alpha, G_0 = N_p(\boldsymbol{\beta} | \boldsymbol{\mu}, \boldsymbol{\Sigma}))$$

and the priors are specified as in the parametric model.

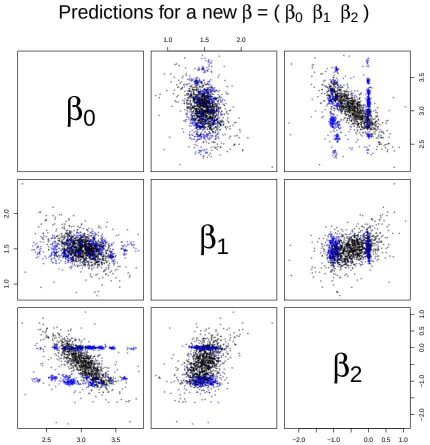
# Fitting the models

The priors all yield conjugate posterior conditionals, so sampling can be done easily with Gibbs

For the DPM, since $p(\cdot)$ is a polynomial and $G_0$ is normal, we can use the Gibbs sampler from Escobar and West.

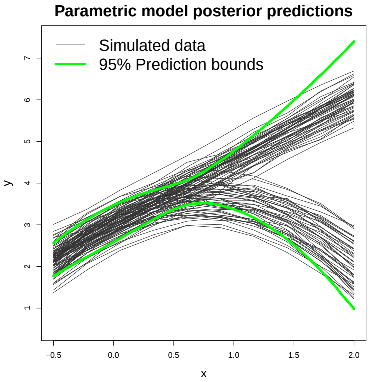
## Simulated data: posterior predictive of a new $\beta_0$

Black – parametric; Blue – DPM)



# Simulated data: posterior predictive of a new $\mathrm{y}_0$

Based on the posterior predictive for $\boldsymbol{\beta}_0$

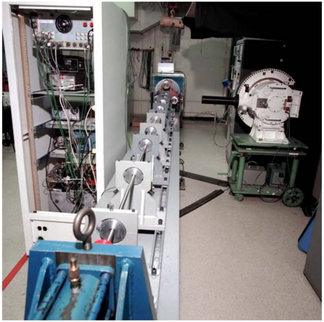


# Hopkinson-bar experiments

A small piece of material is compressed or stretched, either slowly or rapidly

The deformation of the material is measured in two ways: stress and strain

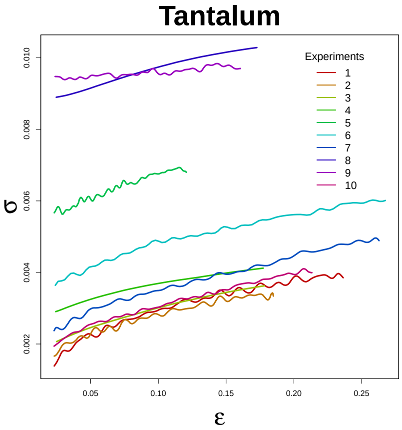
Experiments are done at various temperatures and strain rates



# Hopkinson-bar experiments

We have curves from $n = 10$ experiments

Materials models are typically used to model the data



## Material model

For $\mathbf{x} = (\boldsymbol{\epsilon}_p, \dot{\epsilon}^*, T^*)$ and $\boldsymbol{\theta} = (A, B, n, C, m)$ the Johnson-Cook model is given by

$$h(\mathbf{x}, \boldsymbol{\theta}) = (\mathbf{A} + \mathbf{B}\boldsymbol{\epsilon}_p^{\mathbf{n}})(1 + \mathbf{C}\log \dot{\epsilon}^*)(1 - T^{*\mathbf{m}})$$

where $\dot{\epsilon}^*$ is the experimental strain rate, $T^*$ is the experimental temperature, scaled by melting temperature (in Kelvin), and $\epsilon_p$ is the plastic strain ($x$-axis)

There are other, much better models, but the Johnson-Cook is very simple and is a useful starting point

## Statistical models

Replace $p(\mathbf{x}_i, \boldsymbol{\beta}_i)$ with the Johnson-Cook model $h(\mathbf{x}_i, \boldsymbol{\theta}_i)$ for the parametric and semi-parametric models:

Parametric

$$f(\mathbf{y}_i|\mathbf{x}_i, \boldsymbol{\theta}_i, \tau^2) = N_{k_i}(\mathbf{y}_i|h(\mathbf{x}_i, \boldsymbol{\theta}_i), \tau^2\mathbf{I}_{k_i})$$

DPM

$$f(\mathbf{y}_i|G, \mathbf{x}_i, \tau^2) = \int N_{k_i}(\mathbf{y}_i|h(\mathbf{x}_i, \boldsymbol{\theta}), \tau^2\mathbf{I}_{k_i})dG(\boldsymbol{\theta})$$

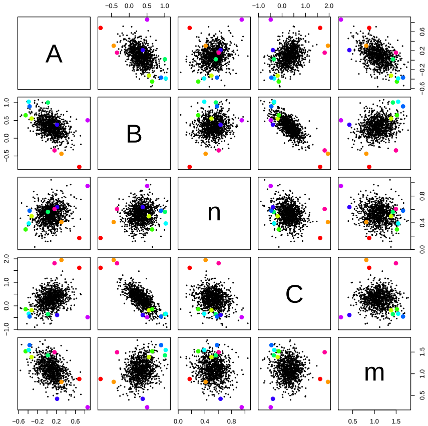
## Fitting the models

Since we no longer have conjugacy with $\boldsymbol{\theta}$ so we update the latent variables using a Metropolis step (Algorithm 6 from Neal (2000))

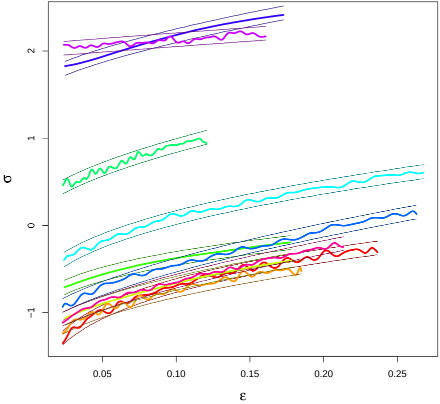
All other priors are the same as before and can be updated with Gibbs steps

# Posteriors for $\theta$

Black – $\boldsymbol{\mu}$; colored dots – means of each $\boldsymbol{\theta}_i$



# Predictions (given the random effects)



## Conclusions and future work

Predictions based a new $\boldsymbol{\theta}_0$ (not shown) were very poor

For covariates $\mathbf{x}^*$ draw a new $\boldsymbol{\theta}_0$ that are similar to one of the $\boldsymbol{\theta}_i$'s

Hierarchical DPs? Nested DPs?