

Dirichlet process mixture model on Hopkinson-bar experiments

Mickey Warner
AMS 241

1 Introduction

Hopkinson-bar (or Split-Hopkinson pressure bar) experiments are compression tests performed on small (e.g. 6.35mm diameter) samples of materials. The material undergoes a deformation which is measured by a strain-stress (ϵ - σ) response. Strain measures changes in the length of the material, and stress measures changes in the cross-sectional area. A single experiment yields a strain-stress curve.

An experiment may fall within two classes of tests: quasi-static or dynamic. In quasi-static tests, the compression occurs very slowly, attempting to measure infinitesimal changes in the response. These are associated with low strain rates, $\dot{\epsilon}$. Dynamic tests occur at high strain rates and are characterized by higher volatility in the response curve.

Other variations of this experiment exist, but the data we will look at in this paper come from either quasi-static or dynamic tests. The data are of Hopkinson-bar tests performed on samples of tantalum by Chen and Gray (1996) (see figure 1), at various strain rates and initial temperatures.

The goal is to model these response curves so we can make accurate predictions for future experiments. The curves themselves (and consequently any predicted curve) may be fed into more complicated materials science models such as those involving 3-D printing. Materials models have been developed to capture the physics and behavior of Hopkinson-bar experiments. We will use a simplified version of a particular model, the Johnson-Cook model, in conjunction with a Dirichlet process mixture model, as our statistical model.

In section 2, we discuss some of the ideas used in modeling these kinds of experiments, from both materials science and statistics perspectives. We discuss Dirichlet process mixture models in section 3 and provide an example of its usage to a simulated data set in section 4. In section 5 we analyze the Hopkinson-bar tests from Chen and Gray (1996). We conclude in section 6 with a discussion.

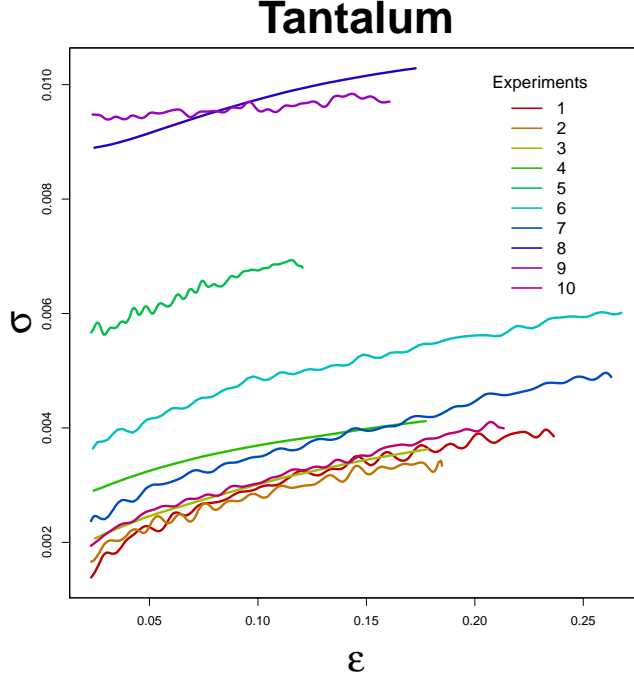


Figure 1: Strain-stress response from ten experiments at various strain rates and temperatures.

2 Brief review of current approaches

Perhaps the simplest materials model that attempts to model the stress σ is the Johnson-Cook model:

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

where \mathbf{x} encodes the plastic strain ϵ_p (the x -axis in figure 1), strain rate $\dot{\epsilon}$, and initial temperature T for a specific experiment, and $\boldsymbol{\theta} = (A, B, n, C, m)$ is the vector of parameters to be estimated. Other models exist that may have more parameters, and hence better predictive capabilities, but model (1) provides us with a convenient starting point.

A typical approach in the materials science community is to assume a common $\boldsymbol{\theta}$ among all experiments and then minimize some goodness-of-fit function. The minimization may be done by optimizing one parameter at a time, holding some parameters fixed (for more complicated models), or even by looking at a graph. We would at least prefer a systematic method of estimating the parameters.

A challenge for any materials model is the potential lack in accurately accounting for all of the necessary physics in the experimental procedure. If a key component is missing, predictions made at untried experimental settings may fail. If too many terms are added, we run the risk of overfitting.

Fugate et al. (2005) fit a hierarchical Bayesian model to Hopkinson-bar experiments where the mean curve is given by a materials model such as (1). By assuming each experiment has its own θ_i , the authors found that the added flexible increased the predictive power of the statistical model. A model with common θ yielded predictive uncertainty that was too conservative. In the hierarchical setting, the mean predictive curve for a new experiment was more accurate, but the predictive variance necessarily increased.

A parametric assumption made by Fugate et al. (2005) is that the population distribution for the θ_i 's is (truncated) multivariate normal. For data sets or materials models where the random effects θ_i do not come from a normal distribution, predictions based on this hierarchical model do very poorly.

3 Dirichlet process mixture models

Before introducing the Dirichlet process mixture (DPM) model, we write out the hierarchical model used by Fugate et al. (2005). For experiments $i = 1, \dots, n$, we have a vector of measured stresses, denoted \mathbf{y}_i , modeled as

$$\mathbf{y}_i | \mathbf{x}_i, \theta_i \stackrel{ind}{\sim} N_{k_i}(\mathbf{y}_i | \boldsymbol{\sigma}(\mathbf{x}_i, \theta_i), \tau^2 \mathbf{I}_{k_i}) \quad (2)$$

$$\theta_i | \boldsymbol{\mu}, \boldsymbol{\Sigma} \stackrel{iid}{\sim} N_p(\theta_i | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \times \mathbb{1}(\theta_i \in C) \quad (3)$$

$$\boldsymbol{\mu} \sim N_p(\boldsymbol{\mu} | \mathbf{m}, \mathbf{S}) \times \mathbb{1}(\boldsymbol{\mu} \in C)$$

$$\boldsymbol{\Sigma} \sim IW(\boldsymbol{\Sigma} | \mathbf{V}, d)$$

$$\tau^2 \sim IG(\tau^2 | a_{\tau^2}, b_{\tau^2})$$

where $\mathbf{x}_i = (\boldsymbol{\epsilon}_{pi}, \dot{\epsilon}_i, T_i)$ is a contains the plastic strains $\boldsymbol{\epsilon}_{pi}$, strain rate $\dot{\epsilon}_i$, and initial temperature T_i for each i , $\boldsymbol{\sigma}(\cdot)$ is given in eq. (1), $\mathbf{m}, \mathbf{S}, \mathbf{V}, d, a_{\tau^2}, b_{\tau^2}$ are known, and $C \subseteq \mathbb{R}^p$ is a set of constraints specified by the materials model. The function $\mathbb{1}(x \in A)$ returns 1 if x is an element of the set A and 0 otherwise. We will call this model the parametric model.

Depending on the materials model, the normal distribution assumption (3) for the random effects may be inappropriate. Therefore, we may wish to be more flexible than this specification.

The Dirichlet process mixture model assumes a non-parametric prior on

the distribution of random effects. We replace (3) with

$$\begin{aligned}\boldsymbol{\theta}_i|G &\stackrel{iid}{\sim} G \\ G|\alpha, \boldsymbol{\mu}, \boldsymbol{\Sigma} &\sim DP(\alpha, G_0 = N_p(\boldsymbol{\theta}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \times \mathbb{1}(\boldsymbol{\theta}_i \in C))\end{aligned}$$

and keep all else from the parametric model the same.

The posterior distributions for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and τ^2 in both models can be updating using Gibbs steps. Since $\boldsymbol{\sigma}(\cdot)$ is non-linear, updating $\boldsymbol{\theta}_i$ in both models is made with Metropolis steps. In the parametric model, the step should be straight-forward. In the DPM model, we update $\boldsymbol{\theta}_i$ after integrating out the infinite-dimensional parameter G . Algorithm 6 by Neal (2000) is used for these updates.

4 Simulation example

To illustrate some practical differences between the parametric and the DPM models, we consider a toy example (see figure 2). We generate $i = 1, \dots, n = 100$ curves, each denoted by \mathbf{y}_i , with random intercepts and linear and quadratic coefficients, according to

$$\begin{aligned}\mathbf{y}_i &= \beta_{i0}\mathbf{1}_{k_i} + \beta_{i1}\mathbf{x}_i + \beta_{i2}\mathbf{x}_i^2 + \boldsymbol{\epsilon}_i \\ \boldsymbol{\epsilon}_i &\stackrel{iid}{\sim} N(0, 0.05^2\mathbf{I}_{k_i}) \\ \beta_{i0} &\stackrel{iid}{\sim} N(3, 0.3^2) \\ \beta_{i1} &\stackrel{iid}{\sim} N(1.5, 0.1^2) \\ \beta_{i2} &\stackrel{iid}{\sim} 0.5\delta_0(\cdot) + 0.5N(-1, 0.1^2)\end{aligned}$$

where $\mathbf{x}_1 = \dots = \mathbf{x}_n = (-0.5, \dots, 2)^\top$ are vectors of $k_i = 10$ equally spaced points and \mathbf{x}^2 denotes element-wise squaring. The vector \mathbf{x}_i should not be thought of only as covariates, but rather the locations where measurements were made, say for repeated measures or functional data.

We fit the simulated data to both the parametric and DPM models, where $\boldsymbol{\sigma}(\cdot)$ is replaced with the polynomial $p(\mathbf{x}, \boldsymbol{\beta}) = \beta_0\mathbf{1}_{k_i} + \beta_1\mathbf{x} + \beta_2\mathbf{x}^2$. So we are assuming that every coefficient in the linear model is a random effect.

Posterior predictions for each model are shown in figures 3 and 4. Even though both models estimated the random effects $\boldsymbol{\beta}_i$ for each observation to be nearly the same, we can see a clear difference in the predictions. Figure 3 shows a pairs plot for the predictive distribution for a new random effect $\boldsymbol{\beta}_0$, and 4 shows a 95% predictive region for a new observation \mathbf{y}_0 . Notice

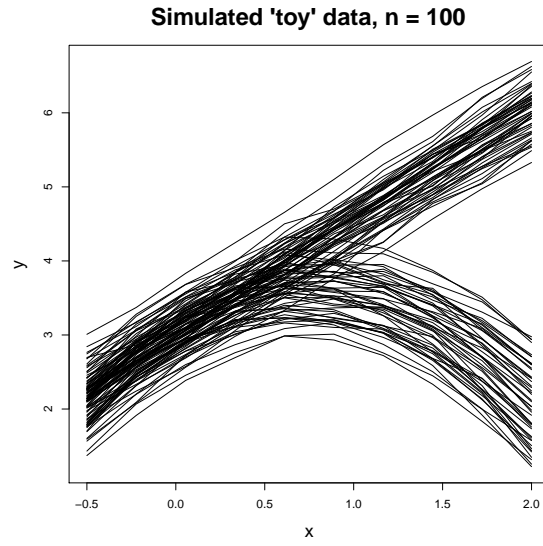


Figure 2: Simulated data set.

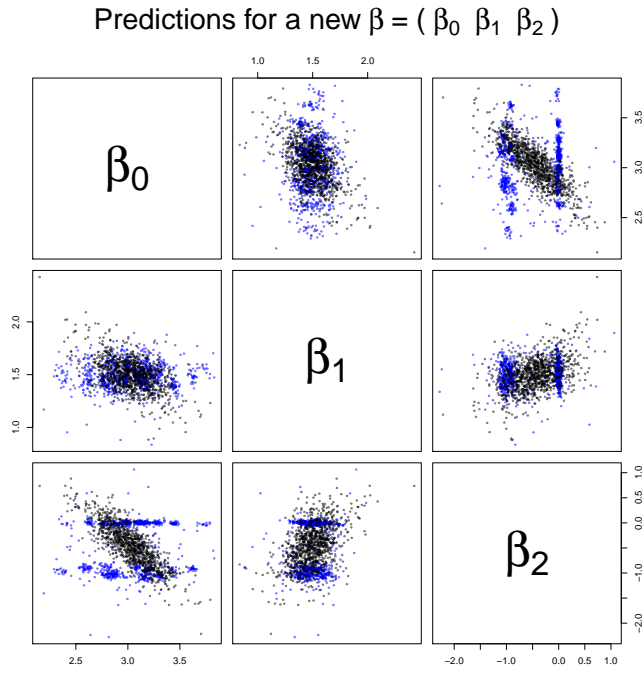


Figure 3: Posterior predictive distribution of a new random effect β_0 . Black: parametric model. Blue: DPM model.

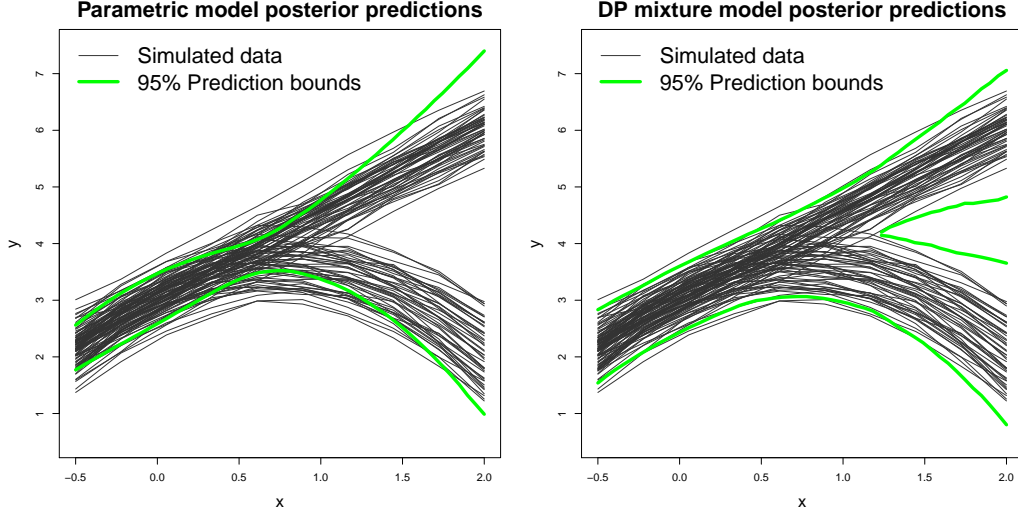


Figure 4: The posterior predictive distribution of a new observation \mathbf{y}_0 based on β_0 shown with the green 95% prediction bounds. Left: parametric model. Right: DPM model.

that the parametric model produces a normal distribution for β_0 , while the DPM more correctly captures the bimodality in the quadratic term. In figure 4 we observe that the parametric model is unable to capture the distinction between the lines and the parabolas in the data. However, the DPM model is capable of doing just that since the non-parametric prior allows for flexible random effects distributions.

5 Analysis of the Hopkinson-bar experiments

We now turn our attention back to the Hopkinson-bar experiments of Chen and Gray (1996) (see figure 1). A previous analysis fit the parametric modeled described in section 3 and used the Johnson-Cook model (1) as the mean of the curves. It was noted that the parametric model failed to produce adequate predictions: it was clear that the population distribution for the random effects was not normal.

The Dirichlet process mixture model of section 3 was subsequently fit to the data, with the hope that the non-parametric assumption would yield better results. To our dismay, poor predictions were also produced by this model. Figure 5 shows bivariate projections of the distribution for the baseline mean μ (in black) and the means of the random effects θ_i (in colors). If we know the random effect for a particular experiment, we can achieve very accurate

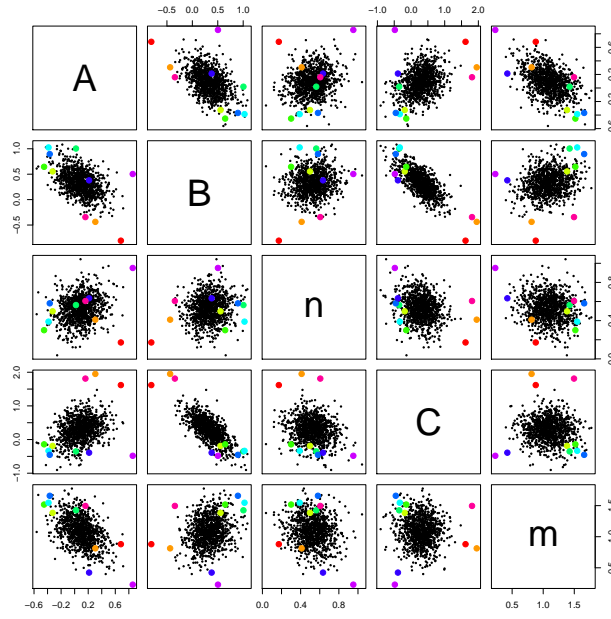


Figure 5: Each two-way dot plots of baseline mean μ (black dots) and the means of each random effect (colored dots).

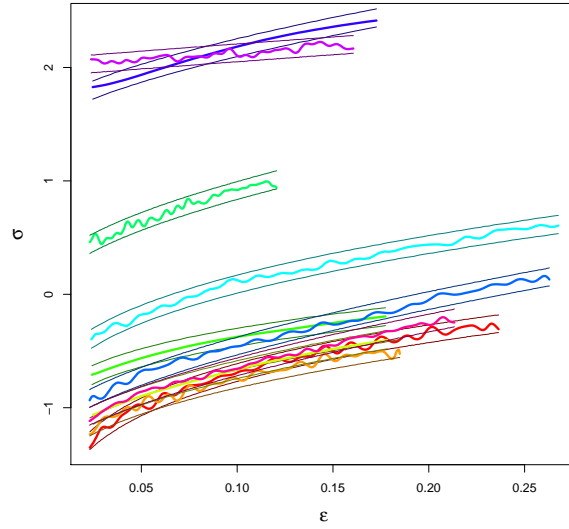


Figure 6: 95% prediction bounds for each observation using its estimated random effect distribution.

predictions (as in figure 6). However, these are meaningless in predicting a new experiment. Predictions based on a new θ_0 were so outrageously inaccurate, with such high variance, that showing them would be pointless.

Though the DPM model allows for a more flexible population distribution, the problem with predictions are due to the Johnson-Cook model being very sensitive to a change in parameters. Consider again figure 5. The colored dots correspond to the mean of posterior random effects for each observation. The figure does not show it, but the variance around each of these points is very small. So when predicting a new experiment, with a given strain rate and temperature, the “correct” random effect to choose from will only be selected a few times and anything away from this “correct” value would produce highly unrealistic predictions. In this sense, the random effects are *too* unique across experiments and do not share enough information.

6 Discussion

In this paper we fit a Dirichlet process mixture model to Hopkinson-bar experiments. At particular strain rates and temperatures, these experiments result in strain-stress curves. Materials models such as the Johnson-Cook model (1) attempt to model the curves. Early attempts at fitting the materials models showed that a common parameter vector may not be optimal, hence a hierarchical approach has been taken.

Depending on the choice of materials model, the multivariate normal assumption on the population distribution may be suspect. To address this, we assumed a non-parametric prior on the population distribution. However, similar difficulties existed in both the parametric setting and the non-parametric setting. Predicting new experiments still faced the challenge of an overly sensitive materials model.

Since the Johnson-Cook model is known to be overly simplistic, we do not expect this kind of poor predictive behavior to be found when using another model such as mechanical threshold stress. The goal here was to identify potential weaknesses in the DPM model and in the materials model. Had the materials model been correctly specified, when making predictions we would have drawn θ s too frequently which yield unrealistic results. A possible next step could be to use a new covariate \mathbf{x}^* to decide θ_i to draw from. I suspect that if \mathbf{x}^* is near the \mathbf{x}_i of an experiment used to fit the model, drawing from that experiments θ_i may produce better results. Perhaps this may be done by making μ a function of the covariates.

References

- Chen, S. R. and Gray, G. T. (1996), “Constitutive behavior of tantalum and tantalum-tungsten alloys,” *Metallurgical and Materials Transactions A*, 27, 2994–3006.
- Fugate, M., Williams, B., Higdon, D., Hanson, K. M., Gattiker, J., Chen, S.-R., and Unal, C. (2005), “Hierarchical Bayesian Analysis and the Preston-Tonks-Wallace Model,” *Las Alamos National Lab*.
- Neal, R. M. (2000), “Markov chain sampling methods for Dirichlet process mixture models,” *Journal of computational and graphical statistics*, 9, 249–265.