Vettori di misure di probabilitá aleatorie: dalle copule di Lévy alle misure aleatorie composte

Fabrizio Leisen

University of Kent

http://sites.google.com/site/fabrizioleisen/

Universitá di Modena e Reggio Emilia– June, 9 2015

Outline

- Random probability measures and priors for Bayesian inference
- Dirichlet Process Mixture Models and Vectors of random densities
- Bivariate completely random measures and Lévy copulas
- Compound Random Measures

Joint works with

Antonio Lijoi (U. di Pavia), Jim Griffin (U. of Kent), Dario Spanó (U. of Warwick) and Weixuan Zhu (U. Carlos III de Madrid)

My research is funded by FP7 Marie Curie Career Integration Grant "**Flexible Bayesian Non-parametric priors**", 2014-2018.

References

- Vectors of two parameters Poisson Dirichlet Processes (with A. Lijoi). The journal of Multivariate Analysis, 2011, 102, 482-495.
- A vector of Dirichlet Processes (with A. Lijoi and D. Spanó). Electronic Journal of Statistics, 2013, Vol. 7, 62-90.
- A Multivariate Extension of a Vector of Poisson-Dirichlet Processes (with W. Zhu). Journal of Nonparametric Statistics, Vol 27, 89-105, 2015.
- Compound Normalized Random Measures (with J. Griffin). Under revision.

Random probabilities and Exchangeability

$X^{(\infty)} = (X_n)_{n\geq 1}$	sequence of $\mathbb{X}$–valued observables defined on $(\Omega, \mathscr{F}, \mathbb{P})$
$\boldsymbol{P}_{\mathbb{X}}$	space of probability measures on $(\mathbb{X}, \mathscr{X})$

Assumption. $X^{(\infty)}$ is exchangeable: for any $n \geq 1$ and permutation π of $(1, \ldots, n)$

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{\pi(1)}, \ldots, X_{\pi(n)})$$

de Finetti's representation theorem. The sequence $X^{(\infty)}$ is exchangeable if and only if

$$\mathbb{P}\left[X^{(\infty)} \in A\right] = \int_{\boldsymbol{P}_{\mathbb{X}}} p^{\infty}(A)\, Q(\mathrm{d}p) \qquad \forall A \in \mathscr{X}^{\infty}$$

where p^{∞} is the infinite product measure $p \times p \times \cdots$ on $(\mathbb{X}^{\infty}, \mathscr{X}^{\infty})$

Random Probability measure: Random element with values in $\boldsymbol{P}_{\mathbb{X}}$

Conditional Indipendence and Bayesian inference

Conditional independence: given a random probability measure $\tilde{p}$ with distribution Q

$$\Pr\left[X_1 \in A_1, \ldots, X_n \in A_n \mid \tilde{p}\right] = \prod_{i=1}^{n} \tilde{p}(A_i) \qquad \forall A_i \in \mathscr{X}$$

Q acts as a prior for Bayesian inference: the de Finetti measure of the sequence $X^{(\infty)}$

A model for Bayesian Inference:

For any $n \geq 1$

$$\begin{aligned} X_i \mid \tilde{p} &\overset{\text{iid}}{\sim} \tilde{p} \qquad (i = 1, \ldots, n) \\ \tilde{p} &\sim Q \end{aligned}$$

If Q is degenerate on a subset of $\boldsymbol{P}_{\mathbb{X}}$ indexed by a finite–dimensional parameter, then the inferential problem is said parametric.

Conditional Indipendence and Bayesian inference

Conditional independence: given a random probability measure $\tilde{p}$ with distribution Q

$$\Pr[X_1 \in A_1, \dots, X_n \in A_n \mid \tilde{p}] = \prod_{i=1}^{n} \tilde{p}(A_i) \qquad \forall A_i \in \mathscr{X}$$

Q acts as a prior for Bayesian inference: the de Finetti measure of the sequence $X^{(\infty)}$

A model for Bayesian Inference:

For any $n \geq 1$

$$\begin{aligned} X_i \mid \tilde{p} &\overset{\text{iid}}{\sim} \tilde{p} \qquad (i = 1, \dots, n) \\ \tilde{p} &\sim Q \end{aligned}$$

If Q is degenerate on a subset of $\boldsymbol{P}_{\mathbb{X}}$ indexed by a finite–dimensional parameter, then the inferential problem is said parametric.
Example:

$$\begin{aligned} X_i \mid (\mu, \sigma^2) &\overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2) \quad (i = 1, \dots, n) \\ (\mu, \sigma^2) &\sim Q \end{aligned}$$

In this case Q is a probability distribution for (μ, σ^2) on $\mathbb{R} \times \mathbb{R}^+$

If Q does not degenerate on a subset of $\boldsymbol{P}_{\mathbb{X}}$ depending on a finite–dimensional parameter, then Q defines a nonparametric prior distribution for Bayesian inference

Possible definition of a nonparametric prior Q: probability distribution of

$$\tilde{p} = g(\tilde{\mu})$$

where $\tilde{\mu}$ is a completely random measure (CRM)

Completely random measure

- $\boldsymbol{M}_{\mathbb{X}}$ space of boundedly finite measures on $(\mathbb{X}, \mathscr{X})$

$$\mu \in \boldsymbol{M}_{\mathbb{X}} \implies \mu(A) < \infty \quad \text{for any} \quad A \in \mathscr{X} \quad \text{bounded}$$

- $\mathscr{M}_{\mathbb{X}}$ Borel σ–algebra on $\boldsymbol{M}_{\mathbb{X}}$
- $\tilde{\mu}$ measurable function from $(\Omega, \mathscr{F}, \mathbb{P})$ to $(\boldsymbol{M}_{\mathbb{X}}, \mathscr{M}_{\mathbb{X}})$ such that for any

$$A_1, \ldots, A_n \in \mathcal{X} \text{ with } A_i \cap A_j = \emptyset$$

  for any $i \neq j$, the random variables

$$\tilde{\mu}(A_1), \ldots, \tilde{\mu}(A_n)$$

  are mutually independent.

Then $\tilde{\mu}$ is a *completely random measure* on $(\mathbb{X}, \mathscr{X})$.

See Kingman (1967), Kingman (1993) or Daley & Vere-Jones (2003)

If $\tilde{\mu}$ is a CRM on $\mathbb{X}$, then

$$\tilde{\mu} = \tilde{\mu}_c + \sum_{i=1}^{q} J_i \, \delta_{x_i}$$

where

- $q \in \{0, 1, \ldots\} \cup \{\infty\}$
- the x_i's are the fixed points of discontinuity of $\tilde{\mu}$
- the J_i's are independent non–negative random variables
- $\tilde{\mu}_c$ is a CRM with no fixed discontinuities
- the J_i's and $\tilde{\mu}_c$ are independent

$\tilde{\mu}_c$ is a CRM

- with no fixed jumps
- characterized by a *Lévy intensity* ν

ν measure on $\mathbb{R}^+ \times \mathbb{X}$ such that

(i) integrability condition

$$\int_{\mathbb{X}\times\mathbb{R}^+} \min\{s, 1\}\, \nu(\mathrm{d}x, \mathrm{d}s) < \infty$$

(ii) Laplace functional transform

$$\mathbb{E}\left[\mathrm{e}^{-\tilde{\mu}_c(f)}\right] = \exp\left\{-\int_{\mathbb{R}^+\times\mathbb{X}} [1 - \mathrm{e}^{-sf(x)}]\, \nu(\mathrm{d}x, \mathrm{d}s)\right\}$$

where $f : \mathbb{X} \to \mathbb{R}$ is a measurable function such that $\mathbb{P}[\tilde{\mu}_c(|f|) < \infty] = 1$

In what follows $q = 0$ and $\tilde{\mu} = \tilde{\mu}_c$

CRMs and nonparametric priors

Gamma CRM

- α finite measure on $(\mathbb{X}, \mathscr{X})$
- $\nu(\mathrm{d}x, \mathrm{d}s) = \alpha(\mathrm{d}x)\mathrm{e}^{-s}s^{-1}\,\mathrm{d}s$
- $\tilde{\mu}(A) \sim \text{Gamma}(1, \alpha(A))$

and $G = \tilde{\mu}/\tilde{\mu}(\mathbb{X})$ is a Dirichlet process with baseline measure α (Ferguson, AoS 1973)

σ–stable CRM

- α finite measure on $(\mathbb{X}, \mathscr{X})$
- $\nu(\mathrm{d}x, \mathrm{d}s) = \alpha(\mathrm{d}x)\sigma s^{-1-\sigma}\,\mathrm{d}s/\Gamma(1-\sigma)$ where $\sigma \in (0, 1)$

and $G = \tilde{\mu}/\tilde{\mu}(\mathbb{X})$ is a normalized σ–stable prior (Kingman, JRSSB 1975)

Two parameters PD process

- parameters $\sigma \in (0,1)$ and $\theta > -\sigma$
- $\mathbb{P}_{\boldsymbol{\sigma}}$ probability distribution (p.d.) of the σ–stable CRM
- $\mathbb{P}_{(\boldsymbol{\sigma},\boldsymbol{\theta})}$ p.d. absolutely continuous with respect to $\mathbb{P}_\sigma$ and such that

$$\frac{\mathbb{P}_{(\sigma,\theta)}(\mathrm{d}\mu)}{\mathbb{P}_\sigma(\mathrm{d}\mu)} = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\sigma+1)}\,\mu(\mathbb{X})^{-\theta}$$

- $\tilde{\boldsymbol{\mu}}_{(\boldsymbol{\sigma},\boldsymbol{\theta})}$ random measure with p.d. $\mathbb{P}_{(\sigma,\theta)}$ (not a CRM)

The random probability measure

$$G = \frac{\tilde{\mu}_{(\sigma,\theta)}}{\tilde{\mu}_{(\sigma,\theta)}(\mathbb{X})}$$

is a two parameter Poisson–Dirichlet process PD(σ,θ) (Pitman, PTRF 1995)

normalized $\boldsymbol{\sigma}$–stable $=$ PD$(\boldsymbol{\sigma},\mathbf{0})$

Dirichlet $=$ PD$(\mathbf{0},\boldsymbol{\theta})$

Alternative (stick–breaking) construction of the PD(α, θ), i.e.

$$\tilde{p} = \sum_{j \geq 1} W_j \, \delta_{\varphi_j}$$

- the random variables φ_j are i.i.d. with common p.d. η and $\eta(\{x\}) = 0$ for any $x \in \mathbb{X}$

- the W_j's are independent from the φ_j's

- the weights W_j are determined via a stick–breaking construction, i.e. given a sequence $(V_n)_{n \geq 1}$ of independent random variables with $V_n \sim$Beta$(1 - \alpha, \theta + n\alpha)$

$$W_1 = V_1 \qquad W_j = V_j \prod_{i=1}^{j-1} (1 - V_i) \quad \forall j \geq 2$$

Why CRMs for defining $\tilde{p}$? Representation of its Laplace functional transform which allows to determine

- moments of any functional $\tilde{p}(f)$
- the exchangeable partition probability function (EPPF)

$$\Pi_k^{(n)}(n_1, \ldots, n_k) =$$

$$\Pr[n \text{ data grouped in } k \text{ clusters with frequencies } n_1, \ldots, n_k]$$

- characterization of the posterior distribution of $\tilde{p}$, given $X_1, \ldots, X_n$

Dirichlet process mixture models

A DP mixture (DPM) models(Loh(1988), Hjort (2010)):

- incorporate Dirichlet process priors for parameters in Bayesian hierarchical models of this type: for any $n \geq 1$ and $i = 1, \ldots, n$

$$\begin{aligned} Y_i \mid \theta_i &\overset{\text{ind}}{\sim} \mathcal{K}(\cdot|\theta_i) \\ \theta_i \mid \tilde{p} &\overset{\text{iid}}{\sim} \tilde{p} \\ \tilde{p} &\sim Q \end{aligned}$$

  where $\mathcal{K}$ is a suitable density kernel.

- The model above defines a random density:

$$f(y) = \int \mathcal{K}(y|\theta)\tilde{p}(d\theta) = \sum_{k\geq 1} W_k \mathcal{K}(y|\varphi_k), \qquad (1)$$

- posterior computation, Escobar (1994), Escobar and West (1995).

Groups of observations and a vector of random densities

- Consider a set of observations divided in r sub–samples (or groups):

$$Y_{ij} \qquad i = 1, \dots r, \quad j = 1, \dots, n_i.$$

  Above $Y_{ij} \in \mathbb{Y} \subset \mathbb{R}^d$ is the j-th observation within sub–sample i.
- Typically, the observations of the block i have the same (conditional) density f_i and are (conditionally) independent.
- In non–parametric Bayesian analysis one needs to specify a prior distribution for the vector of densities $(f_1, f_2, \dots, f_r)$.
- Asses a prior for $(f_1, f_2, \dots, f_r)$ and borrow information across blocks.
- Let $\mathcal{K}_i$ be suitable density kernels, then

$$f_i(y) := \int \mathcal{K}_i(y|\theta)\tilde{p}_i(d\theta) \qquad i = 1, \dots, r, \tag{2}$$

  where $(\tilde{p}_1, \dots, \tilde{p}_r)$ is a vector of dependent random probability measures.

Bidimensional CRMs

Goal: define vectors of random probabilities $(\tilde{p}_1, \tilde{p}_2)$ by means of transformations of vectors of dependent CRMs $(\tilde{\mu}_1, \tilde{\mu}_2)$

$(\tilde{\mu}_1, \tilde{\mu}_2)$ is completely random if

$$A \cap B = \varnothing \qquad \Longrightarrow \qquad (\tilde{\mu}_1(A), \tilde{\mu}_2(A)) \perp (\tilde{\mu}_1(B), \tilde{\mu}_2(B))$$

If $(\tilde{\mu}_1, \tilde{\mu}_2)$ does not have fixed discontinuities, then

$$\mathbb{E}\left[\mathrm{e}^{-\lambda_1 \tilde{\mu}_1(A) - \lambda_2 \tilde{\mu}_2(A)}\right] = \exp\left\{-\int_{A \times (\mathbb{R}^+)^2} \left[1 - \mathrm{e}^{-\lambda_1 s_1 - \lambda_2 s_2}\right] \nu(\mathrm{d}x, \mathrm{d}s_1, \mathrm{d}s_2)\right\}$$

for any A in $\mathscr{X}$

Homogeneous case:

$$\nu(\mathrm{d}x, \mathrm{d}s_1, \mathrm{d}s_2) = \alpha(\mathrm{d}x)\, \nu(s_1, s_2)\, \mathrm{d}s_1\, \mathrm{d}s_2$$

Normalization:

$$(\tilde{p}_1, \tilde{p}_2) = \left(\frac{\tilde{\mu}_1}{\tilde{\mu}_1(\mathbb{X})}, \frac{\tilde{\mu}_2}{\tilde{\mu}_2(\mathbb{X})}\right)$$

Given two CRMs $\tilde{\mu}_1$ and $\tilde{\mu}_2$ how can one define $(\tilde{\mu}_1, \tilde{\mu}_2)$ such that it is completely random?

Answer: by working on the Lévy's intensities...

Given two marginal intensities ν_1 and ν_2, how to determine ν such that

$$\int_0^\infty \nu(s_1, s_2)\mathrm{d}s_2 = \nu_1(s_1) \qquad \int_0^\infty \nu(s_1, s_2)\mathrm{d}s_1 = \nu_2(s_2)$$

for any $s_1, s_2 > 0$.

A function $C : [0, \infty]^2 \to [0, \infty]$ such that

(a) $C(s_1, 0) = C(0, s_2) = 0$ for any positive s_1 and s_2

(b) for all $s_1 < t_1$ and $s_2 < t_2$, $C(s_1, s_2) + C(t_1, t_2) - C(s_1, t_2) - C(t_1, s_2) \geq 0$

(c) C has uniform margins, *i.e.* $C(s_1, \infty) = s_1$ and $C(\infty, s_2) = s_2$

is a *positive Lévy copula*.

See Tankov (2003), Cont & Tankov (CRC, 2004, Financial Modeling with jump processes), Kallsen & Tankov (JMVA, 2006), Epifani and Lijoi (Statistica Sinica, 2010).

Examples

Independence copula. For any $(s_1, s_2) \in [0, \infty]^2$, let

$$C_\perp(s_1, s_2) = s_1 \mathbb{1}_{s_2=\infty} + s_2 \mathbb{1}_{s_1=\infty}.$$

Complete dependence copula. Let $\mathbb{X} = \mathbb{R}^+$ and, for any $(s_1, s_2) \in [0, \infty]^2$, set

$$C_\|(s_1, s_2) = \min\{s_1, s_2\}.$$

Clayton copula. For any $\lambda > 0$

$$C_\lambda(s_1, s_2) = \left\{s_1^{-\lambda} + s_2^{-\lambda}\right\}^{-\frac{1}{\lambda}}$$

and λ regulates the degree of dependence between $\tilde{\mu}_1$ and $\tilde{\mu}_2$

$$\lim_{\lambda \to 0} C_\lambda = C_\perp \qquad \text{(independence)}$$

$$\lim_{\lambda \to \infty} C_\lambda = C_\| \qquad \text{(complete dependence)}$$

How determine ν

Marginal tail integrals for ν_1 and ν_2:

$$t \mapsto U_i(t) = \int_t^\infty \nu_i(s)\,\mathrm{d}s \qquad\qquad i = 1, 2$$

Under suitable conditions

$$\nu(s_1, s_2) = \frac{\partial^2}{\partial v_1 \partial v_2} C(v_1, v_2)\Big|_{v_1 = U_1(s_1), v_2 = U_2(s_2)} \nu_1(s_1)\,\nu_2(s_2)$$

See Tankov (2003) and Cont & Tankov (2004).

Vectors of Stable Processes with Clayton copula

Stable marginals

$$\nu_1(s) = \nu_2(s) = \frac{\sigma}{\Gamma(1-\sigma)}\, s^{-1-\sigma}\, \mathbb{1}_{\mathbb{R}^+}(s)$$

with σ in $(0, 1)$ and

$$\nu(s_1, s_2; \lambda) = \frac{(1+\lambda)\sigma^2}{\Gamma(1-\sigma)}\, \frac{(s_1 s_2)^{\sigma\lambda - 1}}{\left\{s_1^{\sigma\lambda} + s_2^{\sigma\lambda}\right\}^{1/\lambda + 2}}\,.$$

In particular if $\lambda = \frac{1}{\sigma}$

$$\nu(s_1, s_2) = \frac{\sigma + 1}{\Gamma(1-\sigma)}\, \frac{\sigma}{(s_1 + s_2)^{\sigma + 2}}$$

This Levy intensity has been used in Leisen and Lijoi (2011) and Zhu and Leisen (2015) to construct a vector of Poisson-Dirichlet processes.

A Vector of Gamma Processes

Gamma marginals

$$\nu_1(s) = \nu_2(s) = s^{-1} e^{-s} \, \mathbb{1}_{\mathbb{R}^+}(s)$$

and

$$\nu(s_1, s_2) = \sum_{i=0}^{1} \frac{1}{(s_1 + s_2)^{i+1}} e^{-(s_1 + s_2)}$$

The Levy Copula behind this process is:

$$C(y_1, y_2) = \Gamma(0, \Gamma^{-1}(0, y_1) + \Gamma^{-1}(0, y_2)) \tag{3}$$

where $\Gamma(a, x) = \int_x^\infty s^{a-1} \, \mathrm{e}^{-s} \, \mathrm{d}s$ is the incomplete gamma function.

From Levy Copulas to Compound Random Measures

If we look at the bivariate stable process defined above from another perspective then

$$\begin{aligned}\nu(s_1, s_2) &= \tfrac{\sigma+1}{\Gamma(1-\sigma)} \tfrac{\sigma}{(s_1+s_2)^{\sigma+2}} \\ &= \tfrac{\sigma+1}{\Gamma(1-\sigma)} \int_0^{+\infty} \tfrac{z^{-\sigma-3}}{\Gamma(\sigma+2)} e^{-(s_1+s_2)/z} dz\end{aligned}$$

In particular,

$$\nu(s_1, s_2) = \int_0^{\infty} z^{-2} \overbrace{e^{-\frac{s_1+s_2}{z}}}^{\prod_{i=1}^{2} f\left(\frac{s_i}{z}\right)} \overbrace{\frac{z^{-\sigma-1}}{\Gamma(\sigma)\Gamma(1-\sigma)}}^{\nu^*(z)} dz$$

where $f(x) = e^{-x}$, $x > 0$ is the exponential distribution (which we will call score distribution) and $\nu^*(z) = \frac{z^{-\sigma-1}}{\Gamma(\sigma)\Gamma(1-\sigma)}$ is a Levy intensity (which we will call directing Levy measure).

Note that,

$$h(s_1, s_2|z) = z^{-2} \prod_{i=1}^{2} f\left(\frac{s_i}{z}\right) \quad s_1, s_2 > 0$$

is a probability density on $\mathbb{R}^+ \times \mathbb{R}^+$

Compound Random Measures

A Compound random measure (CoRM) is a vector of CRMs defined by a *score distribution h* and a *directing Lévy process* ν^* such that

$$\rho_d(ds_1, \ldots, ds_d) = \int h(s_1, \ldots, s_d|z)\, ds_1 \cdots ds_d\, \nu^\star(dz)$$

where $h(\cdot|z)$ is the probability mass function or probability density function of the score distribution with parameters z and $\nu^\star$ is the Lévy intensity of the directing Lévy process which satisfies the condition

$$\int \int \min(1, \| \boldsymbol{s} \|) h(s_1, \ldots, s_d|z)\, d\boldsymbol{s}\, \nu^\star(dz) < \infty$$

where $\| \boldsymbol{s} \|$ is the Euclidean norm of the vector $\boldsymbol{s} = (s_1, \ldots, s_d)$.

Remark: To ensure the existence of the vectors introduced above, the following condition must be satisfied for each $j = 1, \ldots, d$:

$$\nu_j((0, +\infty)) = \int_0^{+\infty} \int h_j(s|z)\nu^\star(dz)ds = +\infty$$

where $h_j(s|z) = \int h(s_1, \ldots, s_{j-1}, s, ds_{j+1}, \ldots, s_d|z)ds_1 \cdots ds_{j-1}ds_{j+1} \cdots ds_d$. If this condition does not hold true, then $\tilde{\mu}_j(\mathbb{X}) = 0$ with positive probability and the normalization does not make sense, see Regazzini, Lijoi and Pruenster (2003).

Compound Random Measures

In Griffin and Leisen (2014), we focus on the sub-class of CoRMs with a continuous score distribution which has independent dimensions and a single scale parameter so that

$$h(s_1,\dots,s_d|z) = z^{-d}\prod_{j=1}^{d} f(s_j/z)$$

where f is a univariate distribution. This implies that each marginal process has the same Lévy intensity of the form

$$\nu_j(ds) = \nu(ds) = \int z^{-1} f(s|z)\, ds\, \nu^\star(dz).$$

In particular, if

$$f(x) = \frac{1}{\Gamma(\phi)} x^{\phi-1}\exp\{-x\}$$

is the Gamma(ϕ, 1) distribution we obtain the following ν^*:

Compound Random Measures

$\nu^*(z)$	Support	Marginal Process
$z^{-1}(1-z)^{\phi-1}$	$0 < z < 1$	Gamma
$z^{-\sigma-1}\frac{\Gamma(\phi)}{\Gamma(\sigma)\Gamma(1-\sigma)}$	$z > 0$	Stable
$\frac{\sigma}{\Gamma(1-\sigma)}z^{-\sigma-1}(1-az)^{\sigma+\phi-1}$	$0 < z < 1/a$	Gen. Gamma

The results are surprising:

- A gamma marginal process arises when the directing Lévy process is a beta process
- A stable marginal process arises when the directing Lévy process is also a stable process
- Generalized gamma marginal processes lead to a directing Lévy process which is a generalization of the beta process (with a power of z which is less than 1) and re-scaled to the interval $(0, 1/a)$.

Compound Random Measures

An intuitive way to introduce the CoRMs is the following: consider the following dependent random probability measures:

$$\tilde{p}_1 = \sum_{i \geq 1} \pi_{1,i} \delta_{X_i} \cdots \tilde{p}_d = \sum_{i \geq 1} \pi_{d,i} \delta_{X_i},$$

where

$$\pi_{j,i} = \frac{m_{j,i} J_i}{\sum_l m_{j,l} J_l}. \tag{4}$$

- The dependence among the $\tilde{p}_j$ is due to the sharing of jumps $(J_i)_{i \geq 1}$, as highlighted in equation (4)
- The $m_{j,i}$'s are the perturbation coefficients that identify specific features of the j-th random measure and they are independent and identically distributed (with score distribution h_j) across the random measures.

Compound Random Measures

This suggests a slice sampling scheme to address posterior inference as in:

J. E. Griffin and S. G. Walker. "Posterior Simulation of Normalized Random Measure Mixtures", Journal of Computational and Graphical Statistics, 20, 241-259.

The posterior distribution can be expressed in a suitable form for MCMC by introducing latent variables $v_1, \ldots, v_d$, $s_{j,i}$ and $u_{j,i}$ for $i = 1, \ldots, n_j$ and $j = 1, \ldots, d$, and integrating over certain jumps. Integrating over these latent variables leads to the correct marginal posterior distribution. A suitable form of posterior distribution for our MCMC method is

$$\prod_{j=1}^{d} v_j^{n_j-1} \left[\prod_{i=1}^{n_j} m^{\dagger}_{j,s_{j,i}} \mathbb{1}\left(u_{j,i} < J_{s_{j,i}}\right) k\left(y_{j,i}|\theta_{s_{j,i}}\right) \right] \exp\left\{ -\sum_{j=1}^{d} v_j \sum_{k=1}^{K} m^{\dagger}_{j,k} J^{\dagger}_k \right\}$$
$$\times \mathbb{E}\left[\exp\left\{ -\sum_{j=1}^{d} v_j \sum_{k=1}^{\infty} m^{\star}_{j,k} J^{\star}_k \right\} \right]$$

where $L = \min_{j=1,\ldots,d; i=1,\ldots,n_g} \{u_{j,i}\}$. The jumps are divided into two disjoint groups $A^{\dagger} = \{(J^{\dagger}_k, m^{\dagger}_{1,k}, \ldots, m^{\dagger}_{d,k}) | J^{\dagger}_k > L\}$ and $A^{\star} = \{(J^{\star}_k, m^{\star}_{1,k}, \ldots, m^{\star}_{d,k}) | J^{\star}_k < L\}$. The set $A^{\dagger}$ has a finite number of elements which is denoted K and $A^{\star}$ has an infinite number of elements.

The full conditional distributions and a general discussion of methods for updating parameters are given in Griffin and Leisen (2014).

fabrizio.leisen@gmail.com F.Leisen@kent.ac.uk

THANK YOU!!!