Supplemental Material for the Paper 'The Kernel Beta Process' of NIPS2011

1 Proof of Theorem 1

We prove (6) is the Lévy measure of the KBP, as defined in Theorem 1, following the same method with which the Lévy measure of the beta process is derived in Theorem 3.1 in [1]. We have the following notation: $A_{n,m}$ is the m^{th} part of the n-equipartition of Ω, $\omega_{n,m}$ is the central point of $A_{n,m}$, $j = \sqrt{-1}$, and $\boldsymbol{u} \in \mathbb{R}^{|\mathcal{S}|}$. The proof proceeds with the following sequence of steps.

$$
\begin{aligned}
\mathbb{E}\{e^{j<\boldsymbol{u},\, \boldsymbol{\mathcal{B}(\mathcal{A})}>}\} &= \mathbb{E}\{e^{\int_{\mathcal{A}} j<\boldsymbol{u},\, \boldsymbol{\mathcal{B}(d\omega)}>}\} \\
&\overset{n\to\infty}{=} \mathbb{E}\{e^{\sum_{A_{n,m}\in\mathcal{A}} j<\boldsymbol{u},\boldsymbol{K_{n,m}}\pi_{n,m}>}\} \\
&= \mathbb{E}\{\Pi_{A_{n,m}\in\mathcal{A}} e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}\pi_{n,m}>}\} \\
&\overset{(a)}{=} \Pi_{A_{n,m}\in\mathcal{A}}\, \mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}}\} \\
&= e^{\sum_{A_{n,m}\in\mathcal{A}} \log \mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}}\}} \\
&\overset{(b)}{=} e^{\sum_{A_{n,m}\in\mathcal{A}}\, \mathbb{E}\{\sum_{k=1}^{\infty} \frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}]^k}{k!}\}} \\
&\overset{(c)}{=} e^{\sum_{A_{n,m}\in\mathcal{A}} \sum_{k=1}^{\infty}\, \mathbb{E}\{\frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>]^k}{k!}\}\mathbb{E}\{[\pi_{n,m}]^k\}} \qquad (1) \\
&\overset{(d)}{=} e^{\sum_{A_{n,m}\in\mathcal{A}} \sum_{k=1}^{\infty}\, \mathbb{E}\{\frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>]^k}{k!}\} \int_0^1 [\pi_{n,m}]^k \cdot \mathrm{Beta}(c(\omega_{n,m})B_0(A_{n,m}), c(\omega_{n,m})(1-B_0(A_{n,m})))d\pi_{n,m}} \\
&= e^{\sum_{A_{n,m}\in\mathcal{A}} \sum_{k=1}^{\infty}\, \mathbb{E}\{\frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>]^k}{k!}\} \frac{[c(\omega_{n,m})B_0(A_{n,m})+k-1]\cdots[c(\omega_{n,m})B_0(A_{n,m})+1]}{[c(\omega_{n,m})+k-1]\cdots[c(\omega_{n,m})+1]} B_0(A_{n,m})} \\
&\overset{n\to\infty}{=} e^{\sum_{A_{n,m}\in\mathcal{A}} \sum_{k=1}^{\infty}\, \mathbb{E}\{\frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>]^k}{k!}\} \frac{(k-1)\cdots 2\cdot 1}{[c(\omega)+k-1]\cdots[c(\omega)+1]} B_0(d\omega)} \\
&= e^{\int_{\mathcal{X}\times\Psi\times[0,1]\times\mathcal{A}} \sum_{k=1}^{\infty} \frac{[j<\boldsymbol{u},\boldsymbol{K}>]^k}{k!} \pi^k H(dx^*)Q(d\psi^*)c(\omega)\pi^{-1}(1-\pi)^{c(\omega)-1} d\pi B_0(d\omega)} \\
&= e^{\int_{\mathcal{X}\times\Psi\times[0,1]\times\mathcal{A}} \sum_{k=1}^{\infty} \frac{[j<\boldsymbol{u},\boldsymbol{K}\pi>]^k}{k!} \nu_{\mathcal{X}}(dx^*,d\psi^*,d\pi,d\omega)} \\
&\overset{(e)}{=} e^{\int_{\mathcal{X}\times\Psi\times[0,1]\times\mathcal{A}} (e^{j<\boldsymbol{u},\boldsymbol{K}\pi>}-1)\nu_{\mathcal{X}}(dx^*,d\psi^*,d\pi,d\omega)}
\end{aligned}
$$

where the steps of the proof are justified as follows. (a): $\boldsymbol{K_{n,m}}$ and $\pi_{n,m}$ are independent on disjoint sets $\{A_{n,m}\}_{m=1}^{n}$; (b): $\log(\mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}} - 1\} + 1) = \mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}} - 1\}$ since $\mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}} - 1\}$ is infinitesimal. By Taylor series expansion: $\mathbb{E}\{e^{j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}} - 1\} = \mathbb{E}\{\sum_{k=1}^{\infty} \frac{[j<\boldsymbol{u},\boldsymbol{K_{n,m}}>\pi_{n,m}]^k}{k!}\}$. (c): $\boldsymbol{K_{n,m}}$ and $\pi_{n,m}$ are independent to each other. (d): By the definition of the beta process:

$\pi_{n,m} \sim \text{Beta}(c(\omega_{n,m})B_0(A_{n,m}), c(\omega_{n,m})(1 - B_0(A_{n,m})))$. (e): By Taylor series expansion backward.

2 Properties of the KBP

Let $\mathcal{B}_x$, $\mathcal{B}_{x'}$ $(x \neq x')$ be a draw of KBP at covariate x and x', for $\forall \mathcal{A} \subset \mathcal{F}$:

$$\begin{aligned}
\mathbb{E}\mathcal{B}_x(\mathcal{A}) &= \mathbb{E}\int_{\mathcal{A}} \mathcal{B}_x(d\omega) = \int_{\mathcal{A}} \mathbb{E}\mathcal{B}_x(d\omega) = \int_{\mathcal{A}} \mathbb{E}(K_x B(d\omega)) \\
&= \int_{\mathcal{A}} \mathbb{E}K_x \mathbb{E}B(d\omega) = \int_{\mathcal{A}} B_0(d\omega)\mathbb{E}K_x = B_0(\mathcal{A})\mathbb{E}K_x \\
\text{Cov}(\mathcal{B}_x(\mathcal{A}), \mathcal{B}_{x'}(\mathcal{A})) &= \int_{\mathcal{A}} \text{Cov}(\mathcal{B}_x(d\omega), \mathcal{B}_{x'}(d\omega)) = \int_{\mathcal{A}} \text{Cov}(K_x B(d\omega), K_{x'} B(d\omega)) \qquad (2) \\
&= \int_{\mathcal{A}} \mathbb{E}(K_x K_{x'})\mathbb{E}(B^2(d\omega)) - \mathbb{E}K_x \mathbb{E}K_{x'} B_0^2(d\omega) \\
&= \mathbb{E}(K_x K_{x'}) \int_{\mathcal{A}} \frac{B_0(d\omega)(1 - B_0(d\omega))}{c(\omega) + 1} - \text{Cov}(K_x, K_{x'}) \int_{\mathcal{A}} B_0^2(d\omega)
\end{aligned}$$

We are especially interested in the conditional correlation between $\mathcal{B}_x(\mathcal{A})$ and $\mathcal{B}_{x'}(\mathcal{A})$ when kernel parameters $\{x_i^*, \psi_i^*\}_{i=1}^{\infty}$ are given. Denote the mass parameter $\gamma = B_0(\Omega)$. In practice we truncate the number of terms used in (7) to I, then we have the expectation of π_i: $\alpha = \gamma/I$ with $\pi_i \sim \text{Beta}(c\alpha, c(1-\alpha))$ for $i = 1, 2, \cdots, I$. Denote $\boldsymbol{K_x} = (K(x, x_1^*, \psi_1^*), \cdots, K(x, x_i^*, \psi_i^*), \cdots)^T$ with $i : \omega_i \in \mathcal{A}$. Hence:

$$
\begin{aligned}
\mathrm{Corr}(\mathcal{B}_x(\mathcal{A}),\mathcal{B}_{x'}(\mathcal{A})) &= \frac{\int_{\mathcal{A}}\mathrm{Cov}(K_xB(d\omega),K_{x'}B(d\omega))}{\{\int_{\mathcal{A}}\mathrm{Var}(K_xB(d\omega))\cdot\int_{\mathcal{A}}\mathrm{Var}(K_{x'}B(d\omega))\}^{\frac{1}{2}}}\\
&= \frac{\sum_{i:\omega_i\in\mathcal{A}}K(x,x_i^*,\psi_i^*)K(x',x_i^*,\psi_i^*)\mathrm{Var}(\pi_i)}{\{\sum_{i:\omega_i\in\mathcal{A}}K^2(x,x_i^*,\psi_i^*)\mathrm{Var}(\pi_i)\cdot\sum_{i:\omega_i\in\mathcal{A}}K^2(x',x_i^*,\psi_i^*)\mathrm{Var}(\pi_i)\}^{\frac{1}{2}}}\\
&= \frac{\frac{\alpha(1-\alpha)}{c+1}\sum_{i:\omega_i\in\mathcal{A}}K(x,x_i^*,\psi_i^*)K(x',x_i^*,\psi_i^*)}{\frac{\alpha(1-\alpha)}{c+1}\{\sum_{i:\omega_i\in\mathcal{A}}K^2(x,x_i^*,\psi_i^*)\cdot\sum_{i:\omega_i\in\mathcal{A}}K^2(x',x_i^*,\psi_i^*)\}^{\frac{1}{2}}}\\
&= \frac{\sum_{i:\omega_i\in\mathcal{A}}K(x,x_i^*,\psi_i^*)K(x',x_i^*,\psi_i^*)}{\{\sum_{i:\omega_i\in\mathcal{A}}K^2(x,x_i^*,\psi_i^*)\cdot\sum_{i:\omega_i\in\mathcal{A}}K^2(x',x_i^*,\psi_i^*)\}^{\frac{1}{2}}}\\
&= \frac{<\boldsymbol{K_x},\boldsymbol{K_{x'}}>}{\|\boldsymbol{K_x}\|_2\cdot\|\boldsymbol{K_{x'}}\|_2}
\end{aligned}
\tag{3}
$$

References

[1] N. L. Hjort. Nonparametric Bayes estimators based on beta processes in models for life history data. *Annals of Statistics*, 1990.