Теоретическое задание по вариационному выводу

Солоткий Михаил, 417 группа ВМК МГУ

6 ноября 2018 г.

Пусть $X = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_N\}$, $\boldsymbol{x}_n \in \mathbb{R}^D$ – независимая выборка из смеси распределений Стьюдента

$$p(\boldsymbol{x}) = \sum_{k=1}^{K} w_k \mathcal{T}(\boldsymbol{x}|\boldsymbol{\mu}_k, \Sigma_k, \nu), \quad w_k \geq 0, \ \sum_j w_j = 1. \tag{1}$$

Рассмотрим следующую вероятностную модель со скрытыми переменными:

$$p(X, T, Z|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu) = \prod_{n,k=1}^{N,K} \Big[w_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k, \Sigma_k/z_n) \mathcal{G}(z_n|\nu/2, \nu/2) \Big]^{t_{nk}}. \tag{2}$$

Здесь $t_{nk} \in \{0, 1\}$, $\sum_j t_{nj} = 1$ обозначает принадлежность n-го объекта k-ой компоненте смеси. Очевидно, что неполное правдоподобие $p(X|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu)$ для модели (2) совпадает с правдоподобием выборки X для смеси (1). Поэтому оценки максимального правдоподобия $w_{ML,k}, \mu_{ML,k}, \Sigma_{ML,k}$ для смеси (1) можно искать с помощью вариационного EM-алгоритма для модели (2), в котором на E-шаге апостериорное распределение приближается в семействе

$$q_T(T) q_Z(Z) \approx p(T, Z|X, \boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu).$$

Для выполнения задания требуется:

1. Выписать формулы пересчёта для компонент вариационного приближения $q_T(T)$ и $q_Z(Z)$;
2. Выписать формулы пересчёта параметров w_k, μ_k, Σ_k на M-шаге. Убедиться, что эти формулы переходят в соответствующие формулы с семинара по EM-алгоритму для случая $K = 1$;
3. Расписать функционал $\mathcal{L}(q, \boldsymbol{w}, \boldsymbol{\mu}, \Sigma)$ – нижнюю оценку на $\log p(X|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu)$;
4. Найти формулы для статистик распределений $q_T(T)$ и $q_Z(Z)$, требуемых в предыдущих трёх пунктах.

Решение:

1. После выполнения очередной итерации mean-field аппроксимации распределение q_T останется категориальным (носитель не меняется, оно не сможет стать другим). Т. к. априорное гамма-распределение и нормальная функция правдоподобия сопряжены при переменном множителе в ковариационной матрице, после mean-field аппроксимации q_Z будет гамма-распределением. Функция вероятности априорного распределения на переменные T факторизуется в произведение по объектам: $P(T) = \prod\limits_{n=1}^{N} P(t_n)$; $P(t_n = k) = w_k$

   В результате mean-field аппроксимации $q_T(T)$ по-прежнему будет факторизоваться по объектам (это будет показано дальше). Сейчас выведем формулы для распределения q_Z и покажем, что оно тоже факторизуется по объектам.

   $q_T(T) = \prod\limits_{n=1}^{N} q_n(t_n)$, где t_n - индекс компоненты для n-того объекта;

   $q_Z(Z) = \dfrac{\exp\{\mathbb{E}_{q_T} \ln p(X, T, Z|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu)\}}{\int \exp\{\mathbb{E}_{q_T} \ln p(X, T, Z|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu)\} dZ}$

   $p(X, T, Z|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu) = \prod\limits_{n=1}^{N} w_{t_n} \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_{t_n}, \dfrac{\Sigma_{t_n}}{z_n}) \mathcal{G}(z_n|\frac{\nu}{2}, \frac{\nu}{2})$

   $\ln p(X, T, Z|\boldsymbol{w}, \boldsymbol{\mu}, \Sigma, \nu) = \sum\limits_{n=1}^{N} \ln w_{t_n} + \ln \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_{t_n}, \dfrac{\Sigma_{t_n}}{z_n}) + \ln \mathcal{G}(z_n|\frac{\nu}{2}, \frac{\nu}{2})$

$$\mathbb{E}_{q_T}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu) = \sum_{n=1}^{N}\sum_{k=1}^{K} q_n(t_n=k)\ln w_k - \tfrac{ND}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{K} q_n(t_n=k)\ln\det\Sigma_{t_n} +$$

$$+\tfrac{D}{2}\sum_{n=1}^{N}\ln z_n - \tfrac{1}{2}\sum_{n=1}^{N} z_n \sum_{k=1}^{K} q_n(t_n=k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T\Sigma_k^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_k) + \sum_{n=1}^{N}\ln\mathcal{G}(z_n|\tfrac{\nu}{2},\tfrac{\nu}{2})$$

$$c_1 = \exp\Big\{\sum_{n=1}^{N}\sum_{k=1}^{K} q_n(t_n=k)\ln w_k\Big\}\exp\Big\{-\tfrac{ND}{2}\ln(2\pi)\Big\}\exp\Big\{-\tfrac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{K} q_n(t_n=k)\ln\det\Sigma_{t_n}\Big\}$$

$$c_{n2}(x_n) = \sum_{k=1}^{K} q_n(t_n=k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T\Sigma_k^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_k)$$

$$\exp\Big\{\mathbb{E}_{q_T}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\Big\} = c_1\prod_{n=1}^{N}\Big[z_n^{\frac{D}{2}}\exp\Big\{-\tfrac{1}{2}z_n c_{2n}(x_n)\Big\}\mathcal{G}(z_n|\tfrac{\nu}{2},\tfrac{\nu}{2})\Big] =$$

$$= \left\{\mathcal{G}(z|a,b) = \frac{b^a}{\Gamma(a)}z^{a-1}e^{-bz}\right\} = \prod_{n=1}^{N} c_{3n} z_n^{\frac{D+\nu}{2}-1}\exp\Big\{-\tfrac{1}{2}z_n(c_{2n}(x_n)+\nu)\Big\}$$

$$q_Z(Z) = \prod_{n=1}^{N}\frac{\Big(\frac{1}{2}(c_{2n}(x_n)+\nu)\Big)^{\frac{D+\nu}{2}}}{\Gamma(\frac{D+\nu}{2})} z_n^{\frac{D+\nu}{2}}\exp\Big\{-\tfrac{1}{2}z_n(c_{2n}(x_n)+\nu)\Big\}$$

$$q_T(T) = \frac{\exp\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\}}{\int\exp\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\}dT}$$

$$\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu) = \sum_{n=1}^{N}\Big[\ln w_{t_n} + \ln\mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_{\boldsymbol{t_n}},\frac{\Sigma_{t_n}}{z_n}) + \ln\mathcal{G}(z_n|\tfrac{\nu}{2},\tfrac{\nu}{2})\Big] =$$

$$\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu) = \sum_{n=1}^{N}\ln w_{t_n} - \tfrac{ND}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{n=1}^{N}\ln\det\Sigma_{t_n} + \tfrac{D}{2}\sum_{n=1}^{N}\mathbb{E}_{q_Z}\ln z_n -$$

$$-\tfrac{1}{2}\sum_{n=1}^{N}\mathbb{E}_{q_Z}z_n(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n}) + \tfrac{N\nu}{2}\ln\tfrac{\nu}{2} - N\ln\Gamma(\tfrac{\nu}{2}) + (\tfrac{\nu}{2}-1)\sum_{n=1}^{N}\mathbb{E}_{z_n}\ln z_n - \tfrac{\nu}{2}\sum_{n=1}^{N}\mathbb{E}_{z_n}z_n =$$

$$= \Big\{z_n\sim\mathcal{G}(z_n|a_n,b_n)\Big\} = \sum_{n=1}^{N}\ln w_{t_n} - \tfrac{ND}{2}\ln(2\pi) - \tfrac{1}{2}\sum_{n=1}^{N}\ln\det\Sigma_{t_n} + \tfrac{D}{2}\sum_{n=1}^{N}\left[\frac{\Gamma'(a_n)}{\Gamma(a_n)} - \ln b_n\right] -$$

$$-\tfrac{1}{2}\sum_{n=1}^{N}\frac{a_n}{b_n}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n}) + \tfrac{N\nu}{2}\ln\tfrac{\nu}{2} - N\ln\Gamma(\tfrac{\nu}{2}) + (\tfrac{\nu}{2}-1)\sum_{n=1}^{N}\left[\frac{\Gamma'(a_n)}{\Gamma(a_n)} - \ln b_n\right] - \tfrac{\nu}{2}\sum_{n=1}^{N}\tfrac{a_n}{b_n}$$

$$c_1 = (2\pi)^{\frac{-ND}{2}}\tfrac{\nu}{2}^{\frac{N\nu}{2}}(\Gamma(\tfrac{\nu}{2}))^{-N}\exp\left\{\Big(\tfrac{D+\nu}{2}-1\Big)\sum_{n=1}^{N}\left[\frac{\Gamma(a_n)'}{\Gamma(a_n)} - \ln b_n\right] - \tfrac{\nu}{2}\sum_{n=1}^{N}\tfrac{a_n}{b_n}\right\}$$

$$\exp\Big\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\Big\} = c_1\prod_{n=1}^{N}w_{t_n}\prod_{n=1}^{N}\frac{1}{\sqrt{\det\Sigma_{t_n}}}\prod_{n=1}^{N}\exp\Big\{-\frac{1}{2}\frac{a_n}{b_n}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})\Big\}$$

$$\int\exp\Big\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\Big\}dT = \sum_{t_1\ldots t_n\in\mathcal{D}(t_1\ldots t_n)}\exp\Big\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\Big\} =$$

$$= c_1\sum_{t_1\ldots t_n\in\mathcal{D}(t_1\ldots t_n)}\prod_{n=1}^{N}w_{t_n}\prod_{n=1}^{N}\frac{1}{\sqrt{\det\Sigma_{t_n}}}\prod_{n=1}^{N}\exp\Big\{-\frac{1}{2}\frac{a_n}{b_n}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})\Big\}$$

$$f(t_n) = w_{t_n}\frac{1}{\sqrt{\det\Sigma_{t_n}}}\exp\Big\{-\frac{1}{2}\frac{a_n}{b_n}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})\Big\}$$

$$\sum_{t_1\ldots t_n\in\mathcal{D}(t_1\ldots t_n)}\exp\Big\{\mathbb{E}_{q_Z}\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)\Big\} = \sum_{t_1\ldots t_n\in\mathcal{D}(t_1\ldots t_n)}c_1\prod_{n=1}^{N}f(t_n) =$$

$$= c_1\sum_{t_1\in\mathcal{D}(t_1)}\sum_{t_2\ldots t_n\in\mathcal{D}(t_2\ldots t_n)}f(t_1)\prod_{n=2}^{N}f(t_n) = c_1\sum_{t_2\ldots t_n\in\mathcal{D}(t_2\ldots t_n)}\left(\sum_{t_1\in\mathcal{D}(t_1)}f(t_1)\right)\prod_{n=2}^{N}f(t_n) = c_1\prod_{n=1}^{N}\sum_{t_n\in\mathcal{D}(t_n)}f(t_n)$$

$$q_T(T) = \frac{\prod_{n=1}^{N}f(t_n)}{\prod_{n=1}^{N}\sum_{t_n\in\mathcal{D}(t_n)}f(t_n)}$$

2. $q = q_T q_Z$

$q_T(T) = \prod_{n=1}^{N} q_n(t_n)$, где t_n - индекс компоненты для n-того объекта;

$$\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)=\mathbb{E}_q\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)=\sum_{n=1}^{N}\mathbb{E}_q\Big[\ln w_{t_n}-\tfrac{D}{2}\ln(2\pi)-\tfrac{1}{2}\ln\det\Sigma_{t_n}+\tfrac{D}{2}\ln z_n-$$

$$-\tfrac{1}{2}z_n(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})^T\Sigma_{t_n}^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_{t_n})+\tfrac{\nu}{2}ln\tfrac{\nu}{2}-\ln\Gamma(\tfrac{\nu}{2})+\left(\tfrac{\nu}{2}-1\right)\ln z_n-\tfrac{\nu}{2}z_n\Big]=\sum_{n=1}^{N}\sum_{k=1}^{K}q_n(k)\ln w_k-\tfrac{ND}{2}\ln(2\pi)-$$

$$\tfrac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{K}q_n(k)\ln\det\Sigma_k+\tfrac{D}{2}\sum_{n=1}^{N}\Big[\frac{\Gamma^{'}(a_n)}{\Gamma(a_n)}-\ln b_n\Big]-\tfrac{1}{2}\sum_{n=1}^{N}\tfrac{a_n}{b_n}\sum_{k=1}^{K}q_n(k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T\Sigma_k^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_k)+\tfrac{N\nu}{2}\ln\tfrac{\nu}{2}-$$

$$-N\ln\Gamma(\tfrac{\nu}{2})+\left(\tfrac{\nu}{2}-1\right)\sum_{n=1}^{N}\left[\frac{\Gamma^{'}(a_n)}{\Gamma(a_n)}-\ln b_n\right]-\tfrac{\nu}{2}\sum_{n=1}^{N}\tfrac{a_n}{b_n}$$

$$\begin{cases}\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)\to\max\limits_{\boldsymbol{w}}\\ w_k\geq 0\\ \sum\limits_{k=1}^{K}w_k=1\end{cases}$$

KKT:

$$L(w,\lambda,\mu)=\sum_{n=1}^{N}\sum_{k=1}^{K}q_n(k)\ln w_k+\sum_{k=1}^{K}\lambda_k w_k-\mu\left(\sum_{k=1}^{K}w_k-1\right)$$

$$\nabla_{w_k}L(w,\lambda,\mu)=\frac{\sum\limits_{n=1}^{N}q_n(k)}{w_k}+\lambda_k-\mu$$

$$\begin{cases}\nabla_{w_k}L(w,\lambda,\mu)=0\\ w_k\geq 0\\ \sum\limits_{k=1}^{K}w_k=1\\ \lambda_k\geq 0\\ \lambda_k w_k=0\end{cases}\Leftrightarrow\begin{cases}\mu=\frac{\sum\limits_{n=1}^{N}q_n(k)}{w_k}\\ w_k>0\\ \sum\limits_{k=1}^{K}w_k=1\\ \lambda_k=0\end{cases}\Rightarrow\begin{cases}w_k=\frac{\sum\limits_{n=1}^{N}q_n(k)}{\mu}\\ \sum\limits_{k=1}^{K}w_k=1\end{cases}\Rightarrow\mu=\sum_{n=1}^{N}\sum_{k=1}^{K}q_n(k)=N\Rightarrow$$

$$\Rightarrow w_k=\frac{1}{N}\sum_{n=1}^{N}q_n(k)$$

$$\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)\to\max_{\boldsymbol{\mu}}$$

$$\nabla_{\boldsymbol{\mu}_k}\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)=\sum_{n=1}^{N}\frac{a_n}{b_n}q_n(k)\nabla_{\boldsymbol{\mu}_k}[\boldsymbol{x}_n^T\Sigma_k^{-1}\boldsymbol{x}_n-2\boldsymbol{x}_n^T\Sigma_k^{-1}\boldsymbol{\mu}_k+\boldsymbol{\mu}_k^T\Sigma_k^{-1}\boldsymbol{\mu}_k]=2\sum_{n=1}^{N}\frac{a_n}{b_n}q_n(k)[-\Sigma_k^{-1}\boldsymbol{x}_n+$$

$$+\Sigma_k^{-1}\boldsymbol{\mu}_k]=0\Rightarrow\sum_{n=1}^{N}\frac{a_n}{b_n}q_n(k)\boldsymbol{x}_n=\sum_{n=1}^{N}\frac{a_n}{b_n}q_n(k)\boldsymbol{\mu}_k\Rightarrow\boldsymbol{\mu}_k=\frac{\sum\limits_{n=1}^{N}\frac{a_n}{b_n}q_n(k)\boldsymbol{x}_n}{\sum\limits_{n=1}^{N}\frac{a_n}{b_n}q_n(k)}$$

$$\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)\to\max_{\Sigma}$$

$$\nabla_{\Sigma_k^{-1}}\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)=\frac{1}{2}\sum_{n=1}^{N}q_n(k)\Big[\nabla_{\Sigma_k^{-1}}\ln\det\Sigma_k^{-1}-\frac{a_n}{b_n}\nabla_{\Sigma_k^{-1}}(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T\Sigma_k^{-1}(\boldsymbol{x}_n-\boldsymbol{\mu}_k)\Big]=$$

$$=\frac{1}{2}\sum_{n=1}^{N}q_n(k)\Big[\Sigma_k-\frac{a_n}{b_n}(\boldsymbol{x}_n-\boldsymbol{\mu}_k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T\Big]=0\Rightarrow\Sigma_k=\frac{\sum\limits_{n=1}^{N}\frac{a_n}{b_n}q_n(k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)(\boldsymbol{x}_n-\boldsymbol{\mu}_k)^T}{\sum\limits_{n=1}^{N}q_n(k)}$$

3. $q=q_Tq_Z$

   $q_T(T)=\prod\limits_{n=1}^{N}q_n(t_n)$, где t_n - индекс компоненты для n-того объекта;

   $$\mathcal{L}(q,\boldsymbol{w},\boldsymbol{\mu},\Sigma)=\mathbb{E}_q\ln p(X,T,Z|\boldsymbol{w},\boldsymbol{\mu},\Sigma,\nu)=\sum_{n=1}^{N}\mathbb{E}_q\Big[\ln w_{t_n}-\tfrac{D}{2}\ln(2\pi)-\tfrac{1}{2}\ln\det\Sigma_{t_n}+\tfrac{D}{2}\ln z_n-$$

$$- \tfrac{1}{2} z_n (\boldsymbol{x}_n - \boldsymbol{\mu}_{t_n})^T \Sigma_{t_n}^{-1} (\boldsymbol{x}_n - \boldsymbol{\mu}_{t_n}) + \tfrac{\nu}{2} ln \tfrac{\nu}{2} - \ln \Gamma(\tfrac{\nu}{2}) + \left(\tfrac{\nu}{2} - 1\right) \ln z_n - \tfrac{\nu}{2} z_n \Big] = \sum_{n=1}^{N} \sum_{k=1}^{K} q_n(k) \ln w_k - \tfrac{ND}{2} \ln(2\pi) -$$

$$\tfrac{1}{2} \sum_{n=1}^{N} \sum_{k=1}^{K} q_n(k) \ln \det \Sigma_k + \tfrac{D}{2} \sum_{n=1}^{N} \left[\frac{\Gamma^{'}(a_n)}{\Gamma(a_n)} - \ln b_n \right] - \tfrac{1}{2} \sum_{n=1}^{N} \tfrac{a_n}{b_n} \sum_{k=1}^{K} q_n(k) (\boldsymbol{x}_n - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\boldsymbol{x}_n - \boldsymbol{\mu}_k) + \tfrac{N\nu}{2} \ln \tfrac{\nu}{2} -$$

$$- N \ln \Gamma(\tfrac{\nu}{2}) + \left(\tfrac{\nu}{2} - 1\right) \sum_{n=1}^{N} \left[\frac{\Gamma^{'}(a_n)}{\Gamma(a_n)} - \ln b_n \right] - \tfrac{\nu}{2} \sum_{n=1}^{N} \tfrac{a_n}{b_n}$$

4. $\mathcal{G}(z_n|a_n, b_n) = \dfrac{b_n^{a_n}}{\Gamma(a_n)} z_n^{a_n - 1} e^{-b_n z_n}$

$$\mathbb{E}_{z_n} z_n = \int_0^{+\infty} \frac{b_n^{a_n}}{\Gamma(a_n)} z_n^{a_n} e^{-b_n z_n} dz_n = \int_0^{+\infty} \frac{b_n^{a_n+1} \Gamma(a_n + 1)}{b_n \Gamma(a_n) \Gamma(a_n + 1)} z_n^{a_n} e^{-b_n z_n} dz_n = \frac{a_n}{b_n} \int_0^{+\infty} \frac{b_n^{a_n+1}}{\Gamma(a_n + 1)} z_n^{a_n} e^{-b_n z_n} dz_n = \frac{a_n}{b_n}$$

$$\mathbb{E}_{z_n} \ln z_n = \int_0^{+\infty} \ln z_n \frac{b_n^{a_n}}{\Gamma(a_n)} z_n^{a_n - 1} e^{-b_n z_n} dz_n = \left\{ y_n = b_n z_n, z_n = \frac{y_n}{b_n} \right\} = \int_0^{+\infty} (\ln y_n - \ln b_n) \frac{b_n^{a_n}}{\Gamma(a_n)} \frac{y_n^{a_n - 1}}{b_n^{a_n - 1}} e^{y_n} \frac{1}{b_n} dy_n =$$

$$= \frac{1}{\Gamma(a_n)} \int_o^{+\infty} y_n^{a_n - 1} e^{-y_n} \ln y_n dy_n - \ln b_n = \frac{1}{\Gamma(a_n)} \frac{d}{da_n} \int_0^{+\infty} y_n^{a_n - 1} e^{-y_n} dy_n - \ln b_n = \frac{\Gamma^{'}(a_n)}{\Gamma(a_n)} - \ln b_n$$