arXiv:2603.08925v1 [math.ST] 9 Mar 2026

变分推断中的泛函偏差与切空间几何

Sean Plummer
snplmmr@gmail.com

2026年3月11日

摘要

变分推断通过将贝叶斯后验分布投影到一个易于处理的分布族上进行近似。尽管大多数理论分析采用全局散度度量来评估这种近似的质量，许多应用却依赖特定的后验摘要，例如期望、方差或尾部概率。我们提出一个几何框架，用于分析变分近似下后验泛函的偏差。我们证明，后验泛函的主阶偏差由其在变分族所诱导的变分切空间上的正交分量决定。与该空间对齐的泛函仅产生二阶偏差。对于结构化平均场变分族，我们明确刻画了切空间，并证明其由参数块的分块可加函数构成，而交互分量则决定主阶偏差。在标准的局部渐近正态性条件下，我们进一步推导出后验泛函偏差的显式渐近展开，并证明被忽略的交互方向会导致跨块依赖度量出现一阶畸变。这些结果从几何角度解释了平均场变分推断的若干已知性质，包括对跨块相依的系统性畸变。

1 引言

变分推断（VI）已成为复杂统计模型中进行近似贝叶斯推断的一种广泛使用的方法 [4, 6, 9]。变分方法通过将推断限制在一个易于处理的分布族中，并优化散度准则，为原本在计算上难以处理的后验分布提供可扩展的近似。

尽管变分近似具有计算优势，但它会引入系统性偏差，其结构尚未得到充分理解。变分推断的理论分析大多侧重于衡量近似质量的全局指标，例如 Kullback–Leibler散度或后验分布收缩率 [10, 11, 12, 2, 1]。虽然这些结果为近似的整体表现提供了重要保证，但许多实际应用依赖于特定的后验汇总量，包括期望、方差、协方差或尾部概率。这引出了一个自然的问题：哪些后验泛函可以通过变分近似得到准确估计？

本文从几何视角研究这一问题。将变分解视为后验分布在受限变分族上的 Kullback–Leibler投影，我们证明，后验泛函的主阶偏差由其在由变分族诱导的变分切空间正交方向上的分量决定。位于该空间内的泛函在残差中仅产生二阶偏差，而与其正交的分量则产生

主阶误差。这一分解为给定变分族能够准确表示的后验摘要提供了结构化刻画。

这一分解所依据的几何结构与半参数推断中的经典结果密切相关。在半参数推断中，有效估计量由投影到切空间 [3, 8] 上的正交投影刻画。在变分情形下，切空间由变分族本身决定，而正交补则标识出变分近似未能捕捉的后验变差方向。这一解释自然引出了变分影响函数 的概念，用以刻画后验泛函的主阶偏差。

对于结构化平均场变分族，切空间可以用简单而明确的方式刻画 [6, 9]。具体而言，我们证明，切空间由参数块的分块可加函数构成，而其正交补则对应于耦合多个参数块的交互项。这为平均场近似的若干众所周知的性质提供了直观解释：参数块的加性汇总量能够得到准确表示，而对跨块相依敏感的汇总量则会产生主阶偏差。

我们进一步探究了这种几何结构在正则参数模型中的渐近影响 [8, 7]。当后验分布存在局部高斯近似时，变分近似表现为具有受限协方差结构的高斯投影。这为后验泛函的偏差给出了明确的渐近展开。具体而言，交互敏感泛函（如交叉协方差）在平均场近似下会产生一阶渐近畸变，而与变分切空间对齐的泛函则会自动消除主阶偏差项。

综上，这些结果表明，变分族的几何结构不仅决定了哪些后验摘要能够得到准确表示，也决定了变分近似的偏差在大样本下如何变化。

贡献。 本文的主要贡献如下。

1. 我们为变分Kullback–Leibler投影推导出一种泛函偏差分解，将后验泛函的偏差表示为变分切空间正交补的函数。2. 我们证明，与切空间对齐的后验泛函在变分近似下仅表现出二阶偏差。3. 对于结构化均值场族，我们明确刻画其切空间，并证明该空间由分块可加函数构成，其中交互分量决定主导阶偏差。4. 我们通过涉及后验协方差、线性对比的方差和联合尾概率的示例，展示这一框架。5. 在局部渐近正态性条件下，我们推导出明确的渐近偏差展开，并证明被忽略的交互方向会导致跨块依赖度量发生一阶畸变。

本文其余部分安排如下。第2节介绍变分投影框架，并确立变分切空间的一项关键正交性质。第 3 节推导变分近似的泛函偏差分解。第4节刻画结构化均值场族的切空间，第5节则建立一个显式的两块分解。第6节通过示例阐释该理论的意义。第7节研究偏差分解的局部渐近行为，第8节作总结。

2 变分投影框架

设 $(\Theta,\ \mathcal{A},\ \pi_0)$ 是一个配备参考概率测度 π_0 的可测空间。我们假设后验分布和变分族都关于 π_0 绝对连续。

2.1 后验分布与变分族

设后验分布为

$$\pi(d\theta) = p(\theta)\,\pi_0(d\theta),$$

其中，$p(\theta) > 0$ 表示关于 π_0的后验密度。考虑一个参数变分族

$$\mathcal{Q} = \{q_\lambda(d\theta) = r_\lambda(\theta)\,\pi_0(d\theta) : \lambda \in \Lambda \subset \mathbb{R}^k\}.$$

变分推断通过最小化 Kullback–Leibler散度 [4, 9]来选择一个近似

$$F(\lambda) = \mathrm{KL}(q_\lambda \| \pi) = \int r_\lambda(\theta) \log\left[\frac{r_\lambda(\theta)}{p(\theta)}\right] \pi_0(d\theta).$$

假设存在一个极小值点 $\lambda^* \in \Lambda$，并定义 $q^* = q_{\lambda^*}$。在全文中，我们将 q^*解释为后验分布在变分族上的*KL*投影。关于 q^* 的期望将在希尔伯特空间 $L^2(q^*) = L^2(\Theta, \mathcal{A}, q^*)$中计算。

2.2 残差表示

用于比较 q^* 和 π 的一个便捷量是对数密度残差

$$\Delta(\theta) = \log \frac{q^*(\theta)}{\pi(\theta)}.$$

残差 Δ 衡量变分近似与后验分布之间的对数密度差异。注意，当变分近似高估后验密度时，$\Delta(\theta)$ 为正；低估时则为负。等价地，

$$\frac{\pi(\theta)}{q^*(\theta)} = e^{-\Delta(\theta)}.$$

该表示给出如下恒等式

$$E_\pi[g] = E_{q^*}\left[g\, e^{-\Delta}\right]$$

对于任意 $g \in L^1(\pi)$均成立。该关系将使我们能够关联在 π 和 q^*下的期望。

2.3 切空间

变分族的几何结构可以用其在 q^* 处的切空间来描述。令

$$s_j(\theta) = \partial_{\lambda_j} \log r_\lambda(\theta)\big|_{\lambda=\lambda^*}, \qquad j = 1, \ldots, k,$$

表示变分族的得分函数。变分族在q^* 处的切空间定义为

$$T_{q^*}\mathcal{Q} = \overline{\mathrm{span}\{s_1, \ldots, s_k\}}^{L^2(q^*)}.$$

直观而言，$T_{q^*}\mathcal{Q}$ 描述了变分分布在保持属于变分族的同时可以局部扰动的方向。

2.4 得分正交性

KL 最优性 λ^* 的条件意味着残差 Δ 与切空间正交 .

引理1 (切空间正交性)。 对于所有 $h \in T_{q^*}\mathcal{Q}$,

$$E_{q^*}[h\,\Delta] = 0.$$

证明。 首先注意，对 KL 目标函数求导可得

$$\partial_{\lambda_j} F(\lambda^*) = E_{q^*}[s_j(\Delta + 1)].$$

由于 λ^*是内点极小值， $\partial_{\lambda_j} F(\lambda^*) = 0$。对 q_λ进行归一化可得 $E_{q^*}[s_j] = 0$，因此

$$E_{q^*}[s_j\,\Delta] = 0$$

对于每个 $j = 1$、.....、 k。由于得分函数在 $L^2(q^*)$ 中张成了 $T_{q^*}(\mathcal{Q})$，由内积的线性性和连续性即可得出结论。 □

引理1表明，残差 Δ 位于变分切空间的正交补中。这一几何性质将推动下一节所发展的偏差分解。

3 函数偏差分解

现在，我们将后验期望与变分近似下的期望联系起来。关键观察是，残差

$$\Delta(\theta) = \log \frac{q^*(\theta)}{\pi(\theta)}$$

提供了一种方便的测度变换表示。

假设。 存在 $\eta > 0$ 使得

$$E_{q^*}[e^{\eta|\Delta|}] < \infty.$$

当密度比 $\pi/q*$ 具有次指数尾时，该条件成立。

3.1 测度变换展开

以下引理用残差 Δ 表示后验分布与变分期望之间的差异。

引理2 （测度变换 展开）。令 $g \in L^2(q^*)$ 并定义 $\rho(x) = e^{-x} - 1 + x$。则

$$E_\pi[g] - E_{q^*}[g] = -E_{q^*}[g\Delta] + E_{q^*}[g\,\rho(\Delta)].$$

此外,

$$|E_{q^*}[g\,\rho(\Delta)]| \le C\|g - E_{q^*}[g]\|_{L^2(q^*)}\|\Delta\|^2_{L^2(q^*)}$$

其中常数 C 仅取决于 Δ 的可积性。

证明。 由于 $\pi(\theta)/q^*(\theta) = e^{-\Delta(\theta)}$，我们有

$$E_\pi[g] = E_{q^*}[g\, e^{-\Delta}].$$

利用恒等式 $e^{-x} = 1 - x + \rho(x)$ 可得

$$E_\pi[g] - E_{q^*}[g] = -E_{q^*}[g\Delta] + E_{q^*}[g\rho(\Delta)].$$

余项界由不等式推出

$$|\rho(x)| \le \tfrac{1}{2}x^2 e^{|x|},$$

该不等式由泰勒余项公式以及假设的Δ指数可积性推出。

□

3.2 几何偏差分解

现在，我们将测度变换展开与第2节中确立的切空间正交性结合起来。

定理1（变分投影恒等式）。 设 q^* 表示 π 在变分族 $\mathcal{Q}$ 上的 *KL* 投影。设 $T_{q^*}\mathcal{Q}$ 表示 $\mathcal{Q}$ 在 q^* 处的切空间，且 Δ =对数(q^*/π)。对于任意固定的$g \in L^2(q^*)$，写出正交分解

$$g = g_\parallel + g_\perp \quad with \quad g_\parallel \in T_{q^*}\mathcal{Q}, \qquad g_\perp \perp T_{q^*}\mathcal{Q}.$$

Then

$$E_\pi[g] - E_{q^*}[g] = -\langle g_\perp, \Delta\rangle_{L^2(q^*)} + O\Big(\|\Delta\|^2_{L^2(q^*)}\Big)$$

其中隐常数依赖于 g。

证明。 由引理2可得

$$E_\pi[g] - E_{q^*}[g] = -E_{q^*}[g\Delta] + E_{q^*}[g\rho(\Delta)].$$

写出正交分解

$$g = g_\parallel + g_\perp,$$

其中 $g_\parallel \in T_{q^*}\mathcal{Q}$ 和 $g_\perp \perp T_{q^*}\mathcal{Q}$。由引理1可知，残差 Δ 与切空间正交，因此

$$E_{q^*}[g_\parallel \Delta] = 0.$$

因此

$$E_{q^*}[g\Delta] = E_{q^*}[g_\perp \Delta].$$

将这一恒等式代入引理2中的展开式，得到

$$E_\pi[g] - E_{q^*}[g] = -E_{q^*}[g_\perp\Delta] + E_{q^*}[g\rho(\Delta)].$$

余项有如下界：

$$|E_{q^*}[g\rho(\Delta)]| \le C\|g - E_{q^*}[g]\|_{L^2(q^*)}\|\Delta\|^2_{L^2(q^*)}$$

由引理2可得所述结果。 □

注释 1（投影恒等式）。定理1表明，后验分布泛函的主阶偏差由该泛函中与变分切空间正交的分量决定。具体而言，

$$E_\pi[g] - E_{q^*}[g] = -\langle (I - P_T)g, \Delta \rangle_{L^2(q^*)} + O(\|\Delta\|^2_{L^2(q^*)}),$$

隐常数取决于 g。上文中， P_T 表示到 $T_{q^*}\mathcal{Q}$ 的投影。因此，只有泛函在切空间之外的分量才会对一阶偏差产生贡献。这个量扮演着变分影响分量的角色，类似于经典半参数理论中出现的影响函数。

推论1 （无偏泛函类）。 对于固定的 $g \in T_{q^*}\mathcal{Q}$,

$$E_\pi[g] - E_{q^*}[g] = O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

3.3 凸泛函

对于凸泛函，同样的几何机制可通过次梯度推广。设 T 是定义在 $L^2(q^*)$ 的一个凸子集上的凸泛函，该子集包含 π 和 q^*。回顾一下， $\psi \in \partial T(\pi)$ 表示

$$T(\nu) - T(\pi) \geq \langle \psi, \nu - \pi \rangle \text{ for all } \nu \text{ in the domain of } T,$$

其中 $\langle \psi, \ \nu - \pi \rangle$ 表示自然对偶配对。因此，只要 T 在 π 处存在由次梯度 ψ 表示的一阶展开， T 的主要变化就由线性泛函 $\nu \mapsto \langle \psi, \ \nu - \pi \rangle$ 决定。如果这个次梯度属于变分切空间，那么与定理1中相同的正交性论证就会消去一阶项。

我们继续采用如下约定： π 和 q^* 均视为相对于共同支配测度 π_0的密度。

推论2 (具有次梯度表示的凸泛函**)**。 设 T 是凸子集 $L^2(q^*)$ 上的凸泛函，该凸子集包含 π 和 q^*。假设存在 $\psi \in \partial T(\pi)$ 使得$\psi \in T_{q^*}\mathcal{Q} \cap L^2(q^*)$，并假设 T 在 π 处具有如下形式的一阶展开

$$T(q^*) - T(\pi) = \langle \psi, q^* - \pi \rangle_{L^2(q^*)} + r(q^*, \pi),$$

with

$$r(q^*, \pi) = O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

Then

$$T(q^*) - T(\pi) = O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

证明。 根据假设，

$$T(q^*) - T(\pi) = \langle \psi, q^* - \pi \rangle_{L^2(q^*)} + r(q^*, \pi).$$

配对 $\langle \psi, q^* - \pi \rangle_{L^2(q^*)}$ 是由 ψ生成的线性泛函的偏差，因此应用定理1并取 $g = \psi$ ，得到

$$\langle \psi, q^* - \pi \rangle_{L^2(q^*)} = E_\pi[\psi] - E_{q^*}[\psi] = -\langle (I - P_T)\psi, \Delta \rangle_{L^2(q^*)} + O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

由 $\psi \in T_{q^*}\mathcal{Q}$，可得（$I - P_T$）$\psi = 0$，因此

$$\langle \psi, q^* - \pi \rangle_{L^2(q^*)} = O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

将此结果与假设的 $r(q^*, \pi)$上界结合，可得

$$T(q^*) - T(\pi) = O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

□

4 结构化平均场的几何

现在，我们将前面的分析专门应用于结构化平均场变分族 [6, 4]。在这一设定下，变分切空间可以得到明确刻画，从而清楚地描述哪些后验泛函能够被准确表示。

4.1 结构化平均场族

设参数向量划分为若干块 $\theta = (\theta_{B_1}, \ldots\ldots, \theta_{B_m})$。结构化平均场变分族由以下形式的分布构成:

$$q(\theta) = \prod_{b=1}^{m} q_b(\theta_{B_b}).$$

Let

$$q^*(\theta) = \prod_{b=1}^{m} q_b^*(\theta_{B_b})$$

表示后验分布在该变分族上的 KL投影。

4.2 切向方向

在第2节中，切空间是针对有限维参数变分族定义的，通过与参数 λ 相关的得分函数来刻画。对于结构化平均场族，直接处理完整的乘积类更为自然

$$\mathcal{Q}_{MF} = \Big\{ q(\theta) = \prod_{b=1}^{m} q_b(\theta_{B_b}) \Big\},$$

后者通常是无限维的。因此，在本节中，我们将 q^* 处的切空间定义为 $L^2(q^*)$ 中由经过 q^* 的光滑逐块扰动路径所生成的得分方向构成的闭线性张成。

平均场族的扰动源于各个块因子的扰动。考虑对块 b 进行如下形式的扰动

$$q_{b,\varepsilon}(\theta_{B_b}) = q_b^*(\theta_{B_b}) \frac{e^{\varepsilon f_b(\theta_{B_b})}}{E_{q_b^*}[e^{\varepsilon f_b}]},$$

其中 $f_b \in L^2(q^*_{B_b})$。相应的联合分布扰动为

$$q_\varepsilon(\theta) = q_{b,\varepsilon}(\theta_{B_b}) \prod_{j \neq b} q_j^*(\theta_{B_j}).$$

对数密度求导可得

$$\frac{d}{d\varepsilon} \log q_\varepsilon(\theta)\Big|_{\varepsilon=0} = f_b(\theta_{B_b}) - E_{q_b^*}[f_b].$$

因此，切向 由对块 b 的扰动生成的方向仅取决于变量 θ_B $_b$.

4.3 切空间刻画

现在，我们来刻画结构化平均场族的切空间。

定理 2（结构化平均场的切空间）。令 $T_{q^*}\mathcal{Q}_{MF}$ 表示由结构化平均场族的光滑逐块扰动路径生成的 q^* 处的切空间。那么

$$T_{q^*}\mathcal{Q}_{MF} = \left\{ \sum_{b=1}^{m} f_b(\theta_{B_b}) : f_b \in L_0^2(q^*_{B_b}) \right\},$$

其中 $L_0^2(q^*_{B_b})$ 表示在 $q^*_{B_b}$ 下具有零均值的平方可积函数空间。

证明。根据上述扰动论证，由对块 b 的变化生成的任意切向方向都具有以下形式

$$f_b(\theta_{B_b}) - E_{q_b^*}[f_b].$$

因此，切空间包含所有依赖于单个块的中心化函数。此类扰动的线性组合可得如下形式的函数

$$\sum_{b=1}^{m} f_b(\theta_{B_b})$$

其中每个 f_b 在 $q^*_{B_b}$ 下均为中心化函数。

反过来，根据构造，结构化平均场族中的任意光滑扰动路径都分别通过块因子 $q_b(\theta_{B_b})$ 起作用。因此，其得分方向是仅依赖于各个单独参数块的中心化函数之和，不可能包含涉及多个参数块的交互项。

□

4.4 正交补

切空间的正交补由给定每个参数块时条件期望为零的函数组成。特别地，函数 $g \in L^2(q^*)$ 属于 $T_{q^*}\mathcal{Q}_{MF}^{\perp}$，当且仅当

$$E_{q^*}[g \mid \theta_{B_b}] = 0$$

对于所有参数块 $b = 1,\dots,m$ 均成立。由于 q^* 在各参数块之间可分解，因此与所有关于参数块的加性函数正交，等价于给定每个参数块时条件期望为零。该空间中的函数表示耦合多个参数块的交互分量。

4.5 解释

定理 2 表明，结构化平均场近似能够准确表示参数块的加性泛函。偏差源于泛函中依赖于参数块间交互的分量。结合第 3 节的偏差分解可知，交互项决定了变分近似的主阶偏差。

5 两块分解

为了说明前述结果的含义，我们考虑最简单的非平凡结构化平均场设定，其中包含两个参数块。

5.1 双块均值场

设参数向量被划分为 $\theta = (X, Y)$，并考虑均值场变分族 $q(X, Y) = q_X(X)\, q_Y(Y)$。令 $q^*(X, Y) = q_X^*(X)\, q_Y^*(Y)$ 表示后验分布到该族上的KL投影。根据定理 2，变分切空间由形如 $f(X) + g(Y)$ 的加性函数构成。

5.2 函数分解

任意函数 $h \in L^2(q^*)$ 都可以进行双向Hoeffding（ANOVA）分解 [5]

$$h(X, Y) = \mu + h_X(X) + h_Y(Y) + h_{XY}(X, Y), \tag{1}$$

其中

$$\mu = E_{q^*}[h], \quad h_X(X) = E_{q^*}[h \mid X] - \mu, \quad h_Y(Y) = E_{q^*}[h \mid Y] - \mu,$$

且交互分量满足

$$E_{q^*}[h_{XY} \mid X] = 0, \qquad E_{q^*}[h_{XY} \mid Y] = 0.$$

函数 h_X 和 h_Y 属于变分切空间，而 h_{XY} 属于其正交补。上述分解对应于将 $L^2(q^*)$ 正交分解为加性分量和交互分量。

5.3 偏差分解

将这一表　　示与第 3 节的偏差分解相结合，可得以下结果　　.

命题1。 设 $h \in L^2(q^*)$ 具有方程 (1) 中的分解。则

$$E_\pi[h] - E_{q^*}[h] = -E_{q^*}[h_{XY}\Delta] + O\Big(\|\Delta\|^2_{L^2(q^*)}\Big),$$

其中隐常数依赖于 h。

因此，主阶偏差仅取决于交互分量 h_{XY}。

5.4 解释

结果表明，平均场近似能够准确捕捉参数块的加性汇总，而偏差则源于耦合多个参数块的交互项。这为平均场变分推断的若干著名性质提供了几何解释，包括跨块依赖关系的系统性扭曲。

6 示例

我们通过几个实践中常用的后验摘要，说明前述结果的影响。

6.1 交叉协方差

考虑平均场近似下的两个参数块（X，Y），$q^*(x,\ y) = q_X^*(x)\, q_Y^*(y)$。令 $h(X,\ Y) = u(X)\, v(Y)$ 为平方可积函数 u 和 v。利用第 5 节的分解，

$$h(X,Y) = \mu + h_X(X) + h_Y(Y) + h_{XY}(X,Y),$$

其中

$$h_{XY}(X,Y) = \big(u(X) - E_{q^*}[u(X)]\big)\big(v(Y) - E_{q^*}[v(Y)]\big).$$

函数 h_X 和 h_Y 位于变分切空间 $T_{q^*}\mathcal{Q}$中，而 h_{XY} 位于其正交补中。应用定理1可得

$$E_\pi[u(X)v(Y)] - E_{q^*}[u(X)v(Y)] = -E_{q^*}[h_{XY}\Delta] + O\Big(\|\Delta\|^2_{L^2(q^*)}\Big).$$

因此，主导偏差完全源于交互分量。特别是，跨块协方差对后验分布的交互结构十分敏感。

6.2 线性泛函的方差

设 $\theta = (\theta_1,\ \ldots\ldots,\ \theta_d)$，并考虑线性对比 $L(\theta) = a^\top\theta$。相应的二阶矩泛函为 $h(\theta) = L(\theta)^2$。展开可得

$$h(\theta) = \sum_{i=1}^{d} a_i^2\theta_i^2 + 2\sum_{i<j} a_i a_j \theta_i\theta_j.$$

在完全因子化平均场近似下，第一项求和由依赖于单个坐标的函数构成，因此属于切空间 $T_{q^*}\mathcal{Q}$。第二项求和包含耦合不同坐标的交互项，因此属于正交补。因此，估计 $\mathrm{Var}_\pi(L)$ 时的主导偏差源于这些交互分量。

6.3 联合尾概率

考虑泛函 $F(q) = q\{X > t, Y > s\}$。等价地，这对应于期望 $h(X,Y) = \mathbf{1}\{X > t, Y > s\}$。将此指示函数投影到切空间，对应于超越事件的可加边际分量，而正交分量则表示超出边际预测范围的联合超越超额概率。应用定理1可知，$F(q^*)$ 的主导偏差取决于这一交互分量。这解释了为什么平均场近似可能准确地表示边际尾概率，却会扭曲联合尾行为。

6.4 总结

这些示例说明了一种常见模式：影响函数位于切空间 $T_{q^*}\mathcal{Q}$ 中的后验摘要能够被变分近似准确刻画，而依赖于与切空间正交的分量的摘要则表现出一阶偏差。

7 局部渐近偏差

第3–5节提出的几何分解，用后验泛函位于变分切空间正交补中的分量来刻画其后验泛函的有限样本偏差。在许多统计设定中，后验分布依赖于样本量为 n 的样本，并随着 n增大而集中于真实参数值附近。因此，自然要问几何偏差分解如何与后验分布的经典渐近近似相互作用。

在正则参数模型中，在标准 Bernstein–von Mises 条件 [8] 下，后验分布具有局部高斯近似。因此，应用于此类后验分布的变分近似，其表现类似于对相应局部高斯极限实验进行变分近似。本节的结果分析由此产生的后验泛函渐近偏差，并说明先前识别出的几何结构如何决定主导渐近误差。

令 π_n 表示基于 n 个观测得到的后验分布，并令 q_n^* 表示通过在 $\mathcal{Q}$ 上最小化 $\mathrm{KL}(q\|\pi_n)$ 得到的变分近似。定义残差

$$\Delta_n(\theta) = \log \frac{q_n^*(\theta)}{\pi_n(\theta)}.$$

第 3 节的偏差分解适用于每一对 $(\pi_n,\ q_n^*)$。

命题 2 （渐近投影恒等式）。 令 $g \in L^2(q_n^*)$。记

$$g = g_{\parallel,n} + g_{\perp,n}, \qquad g_{\parallel,n} \in T_{q_n^*}\mathcal{Q}, \quad g_{\perp,n} \perp T_{q_n^*}\mathcal{Q}.$$

Then

$$E_{\pi_n}[g] - E_{q_n^*}[g] = -E_{q_n^*}[g_{\perp,n}\Delta_n] + O\Big(\|\Delta_n\|^2_{L^2(q_n^*)}\Big),$$

其中隐常数取决于 g。

命题 2 的投影恒等式表明，渐近偏差仍由残差 Δ_n 与功能在变分切空间上的正交分量之间的相互作用决定。为了得到明确的偏差表达式，我们现在考虑正则参数模型中出现的局部高斯区域。在此区域中，后验分布和变分近似都以速率 $n^{-1/2}$ 集中，并且可近似为协方差矩阵按 $1/n$缩放的高斯分布。这使得我们能够显式计算主导偏差项。

定理 3 (局部渐近偏差展开)。 假设后验分布和变分近似满足

$$\pi_n \overset{\mathrm{TV}}{\to} N(\mu_n, \Sigma/n), \qquad q_n^* \overset{\mathrm{TV}}{\to} N(\mu_n, V/n)$$

对于固定的正定矩阵 Σ 和 V。设 $g:\ \mathbb{R}^d \to \mathbb{R}$ 在 μ_n 的某个邻域内三次连续可微，且其三阶导数有界。则

$$E_{\pi_n}[g(\theta)] - E_{q_n^*}[g(\theta)] = \frac{1}{2n}\mathrm{tr}\big(H_g(\mu_n)(\Sigma - V)\big) + o_p(n^{-1}).$$

证明概要。 记 $\theta = \mu_n + Z/\sqrt{n}$。在 π_n 下， $Z \sim N(0,\Sigma)$，而在 q_n^* 下， $Z \sim N(0,V)$，误差至多为 $o_p(1)$。对 g 在 μ_n 附近作泰勒展开，得到

$$g(\theta) = g(\mu_n) + \frac{1}{\sqrt{n}}\nabla g(\mu_n)^\top Z + \frac{1}{2n}Z^\top H_g(\mu_n)Z + R_n(Z)$$

其中 $E[R_n(Z)] = o(n^{-1})$。线性项在取期望后消失，二次项则给出

$$E_{\pi_n}[Z^\top H_g Z] - E_{q_n^*}[Z^\top H_g Z] = \mathrm{tr}\big(H_g(\Sigma - V)\big).$$

结论得证。 □

推论3 (平均场交互偏差**)**。 若 q_n^* 是平均场近似，使得 V 为对角矩阵，则主导偏差项取决于海森矩阵$H_g(\mu_n)$中对交互敏感的分量。

推论4. 对于 $g(\theta) = \theta_i\theta_j$ 且 $i \neq j$,

$$E_{\pi_n}[\theta_i\theta_j] - E_{q_n^*}[\theta_i\theta_j] = \frac{\Sigma_{ij}}{n} + o(n^{-1}).$$

定理 3 中的展开表明，主导渐近偏差由后验分布的协方差结构与变分族所能表示的协方差结构之间的不匹配决定。对于平均场近似，这种不匹配恰好出现在变分切空间之外的交互方向上。接下来的结果表明，这一观察与前文建立的几何理论完全吻合：影响方向位于切空间内的泛函会自动消除主导渐近偏差项。

定理4 (切空间泛函的偏差降低**)**。 假设定理 *3* 的条件成立。设 q_n^* 为平均场近似，且 g 为满足 $g \in T_{q_n^*}\mathcal{Q}$的分块可加泛函。则

$$E_{\pi_n}[g] - E_{q_n^*}[g] = o_p(n^{-1}).$$

证明。 根据定理 3,

$$E_{\pi_n}[g] - E_{q_n^*}[g] = \frac{1}{2n}\,\mathrm{tr}(H_g(\Sigma - V)) + o_p(n^{-1}).$$

若 $g \in T_{q_n^*}\mathcal{Q}$，则 g 在局部可表示为若干函数之和，每个函数仅依赖于单个变分块。因此， H_g 不包含任何跨块交互项。对于平均场近似，矩阵 V为对角矩阵，因此 $\Sigma - V$的交互分量与 H_g收缩时会消失。因此

$$\mathrm{tr}(H_g(\Sigma - V)) = 0,$$

由此得到所需结果。 □

定理 3 的一个特别直观的推论是考虑平均场近似下的交叉协方差泛函。

命题 3 (平均场近似下交叉协方差的渐近失真**)**。 假设定理 *3* 的条件成立，并假设 q_n^* 是一个平均场高斯近似，因此其局部协方差矩阵 V 为对角矩阵。令

$$g(\theta) = \theta_i\theta_j, \qquad i \neq j.$$

Then

$$E_{\pi_n}[\theta_i\theta_j] - E_{q_n^*}[\theta_i\theta_j] = \frac{\Sigma_{ij}}{n} + o_p(n^{-1}).$$

特别是，如果 $\Sigma_{ij} \neq 0$，那么在平均场近似下，交叉协方差泛函呈现非消失的一阶渐近偏差。

证明。 对于 $g(\theta) = \theta_i\theta_j$ 且 $i \neq j$，海森矩阵为

$$H_g = e_i e_j^\top + e_j e_i^\top,$$

其中 e_i 表示第 i个标准基向量。应用定理 3 得到

$$E_{\pi_n}[g(\theta)] - E_{q_n^*}[g(\theta)] = \frac{1}{2n}\operatorname{tr}\big(H_g(\Sigma - V)\big) + o_p(n^{-1}).$$

由于 V是对角矩阵，对于 $i \neq j$，$(\Sigma - V)_{ij} = \Sigma_{ij}$。直接计算可得

$$\operatorname{tr}\big(H_g(\Sigma - V)\big) = 2(\Sigma - V)_{ij} = 2\Sigma_{ij},$$

因此

$$E_{\pi_n}[\theta_i\theta_j] - E_{q_n^*}[\theta_i\theta_j] = \frac{\Sigma_{ij}}{n} + o_p(n^{-1}).$$

□

命题 3 表明，在局部高斯情形下，平均场变分推断中被忽略的交互方向会直接导致交叉协方差泛函出现一阶渐近畸变。

8 讨论

我们构建了一个几何框架，用于理解变分近似下后验泛函的偏差。分析表明，偏差由变分切空间 $T_{q^*}\mathcal{Q}$的正交补所决定。与该空间对齐的泛函仅产生二阶偏差，而包含正交于该空间的分量的泛函则会出现一阶误差。该投影恒等式类似于半参数理论中的经典结果：估计误差取决于影响函数中正交于模型切空间[3, 8]的分量。

对于结构化均值场族，切空间可明确刻画为分块可加函数的集合。这带来了一种简单的解释：参数块的加性摘要可由变分近似准确表示，而偏差则源于耦合多个参数块的交互分量。第 5 节中的两块分解说明了这种交互结构如何导致常见后验摘要出现偏差。

8.1 对变分推断的启示

这些结果从几何角度解释了均值场变分推断的若干已知性质。特别是，涉及跨块交互的后验摘要（例如协方差或联合尾概率）取决于与变分切空间正交的方向，因此会出现主阶偏差。相反，各个参数块的加性摘要属于切空间，因而往往能得到准确表示。

这一视角补充了现有的变分推断理论分析，后者侧重于全局散度度量或后验收缩 [10, 11, 12, 1]。此类结果量化了整体近似误差，而本文框架则指出，哪些后验泛函对变分族的结构约束最为敏感。

8.2 结构化变分族

切空间的刻画也阐明了为何结构化变分族往往优于完全因子化近似。扩大块结构会扩展切空间，从而降低其正交补的维数。因此，出现主阶偏差的后验泛函会更少。这一观察为结构化平均场方法的实证成效提供了几何解释。

8.3 未来方向

当前框架的若干扩展可能具有研究价值。首先，可以将该分析应用于更丰富的变分族，包括混合近似或基于归一化流的方法，以理解其几何结构如何影响泛函偏差。其次，变分切空间与坐标上升变分推断算法收敛性质之间的关系值得进一步研究。最后，开发诊断工具以识别接近切空间正交补的后验摘要可能会有所帮助，因为这些量最容易受到变分偏差的影响。

总体而言，结果表明，从后验泛函的角度评估变分近似，可以作为传统基于散度的评估的有益补充。这些结果表明，评估变分族时，不仅应考察全局散度度量，还应考察其切空间所表示的函数类。

AI 致谢

作者使用了大语言模型来协助修改、起草和完善稿件的组织结构。所有数学论证、证明、实验实现和解释均为作者本人完成，并经过独立核验。

参考文献

[1] Pierre Alquier 和 James Ridgway。温度化后验及其变分近似的集中性。《统计学年鉴》，2020 年。[2] Pierre Alquier、James Ridgway 和 Nicolas Chopin。吉布斯后验变分近似的性质。《机器学习研究杂志》，17，2016 年。[3] Peter J. Bickel、Chris A. J. Klaassen、Ya’acov Ritov 和 Jon A. Wellner。《半参数模型的高效与自适应估计》。约翰·霍普金斯大学出版社，1993 年。[4] David M. Blei、Alp Kucukelbir 和 Jon D. McAuliffe。变分推断：统计学家综述。《美国统计学会杂志》，112(518):859–877，2017 年。[5] Wassily Hoeffding。具有渐近正态分布的一类统计量。《数理统计年鉴》，19(3):293–325，1948 年。[6] Michael I. Jordan、Zoubin Ghahramani、Tommi S. Jaakkola 和 Lawrence K. Saul。图模型变分方法导论。《机器学习》，37(2):183–233，1999 年。

[7] Bas J. K. Kleijn 和 Aad W. van der Vaart。模型设定错误下的伯恩斯坦–冯·米塞斯定理。《电子统计学杂志》，6:354–381，2012 年。[8] A. W. van der Vaart。渐近统计学。剑桥大学出版社，2000 年。[9] Martin J. Wainwright 和 Michael I. Jordan。图模型、指数族与变分推断。《机器学习基础与趋势》，1(1–2):1–305，2008 年。[10] Yixin Wang 和 David M Blei。变分贝叶斯的频率学派一致性。《美国统计学会杂志》，114(527):1147–1161，2019 年。[11] Yun Yang、Debdeep Pati 和 Anirban Bhattacharya。 α-变分推断及其统计保证。《统计学年鉴》，48(2):886–905，2020 年。[12] Fengshuo Zhang 和 Chao Gao。变分后验分布的收敛速率。《统计学年鉴》，48(4):2180–2207，2020 年。