Анализ ошибки аппроксимации

30 марта 2021 г.

К. О. Вайсер[1]

Напомним, что работа посвящена эффективному сэмплированию из гауссовских процессов. Стандартная схема с репараметризацией требует минимум $O(m^3)$ операций для сэмплирования m значений. Было разработано несколько методов представления гауссовского процесса для увеличения скорости сэмплирования. Ключевой идеей является разложение апостериорного распределения в сумму двух слагаемых —априорного и некоторого обновления распределения.

$$\underbrace{(f \mid \boldsymbol{u})(\cdot)}_{\text{sparse posterior}} \approx \underbrace{\sum_{i=1}^{\ell} w_i \phi_i(\cdot)}_{\text{weight-space prior}} + \underbrace{\sum_{j=1}^{m} v_j k\left(\cdot, \boldsymbol{z}_j\right)}_{\text{function-space update}},$$

Представление этих слагаемых в определенном виде позволяет увеличить эффективность сэмлирования. Теоретическим обоснованием возможности подобного разложения служит правило Маферона. В данном приложении исследуется величина ошибки аппроксимации при использовании факторизации распределения.

Напомним несколько используемых в теореме утверждений:

Утверждение 1. *(Правило Маферона). Пусть* $\mathbf{a}$ *и* $\mathbf{b}$ *имеют совместное нормальное многомерное распределение. Тогда условное распределение на* $\mathbf{a}$ *при* $\mathbf{b} = \beta$ *вычисляется как*

$$(\boldsymbol{a} \mid \boldsymbol{b} = \boldsymbol{\beta}) \stackrel{\mathrm{d}}{=} \boldsymbol{a} + \operatorname{Cov}(\boldsymbol{a}, \boldsymbol{b}) \operatorname{Cov}(\boldsymbol{b}, \boldsymbol{b})^{-1}(\boldsymbol{\beta} - \boldsymbol{b})$$

Применительно к рассматриваемой задаче это утверждение приобретает следующий вид : пусть $f \sim GP(0, k)$, $\mathbf{Z}$ *- множество точек размера* m *и* $f_m = f(\mathbf{Z})$*. Тогда*

$$\underbrace{(f \mid f_m = \boldsymbol{u})(\cdot)}_{\textit{posterior}} \stackrel{\mathrm{d}}{=} \underbrace{f(\cdot)}_{\textit{prior}} + \underbrace{k(\cdot, \mathbf{Z}) \mathbf{K}_{m,m}^{-1} (\boldsymbol{u} - \boldsymbol{f}_m)}_{\textit{update}}$$

Утверждение 2. *(Неравенство Гёльдера). Пусть* $f \in L^p$*, а* $g \in L^q$*, где* $p, q \geq 1, \frac{1}{p} + \frac{1}{q} = 1$*. Тогда* $f \cdot g \in L^1$*, и* $\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$

[1]Московский физико-технический институт, vajser.ko@phystech.edu

Обратимся теперь к исследованию величины ошибки.

Теорема 1. *Пусть $\mathcal{X} \subseteq \mathbb{R}^d$ —компакт и $f \sim GP(0,k)$- стохастически непрерывный стационарный гауссовский процесс. Обозначим за $f \mid \mathbf{y}$ точное апостериорное распределение, $f^{(s)}$ - апостериорное распределение в терминах функционального подхода, $f^{(d)}$ - апостериорное распределение в терминах объединенного подхода и $f^{(w)}$ - априорное распределение в терминах базисного подхода. Тогда*

$$W_{2,L^2(\mathcal{X})}\left(f^{(d)}, f \mid \boldsymbol{y}\right) \leq \underbrace{W_{2,L^2(\mathcal{X})}\left(f^{(s)}, f \mid \boldsymbol{y}\right)}_{\text{ошибка в апостериорном распределении}} + \underbrace{C_1 W_{2,C(\mathcal{X})}\left(f^{(w)}, f\right),}_{\text{ошибка в априорном распределении}} \quad ,$$

где $W_{2,L^2(\mathcal{X})}, W_{2,C(\mathcal{X})}$ - расстояния Вассерштейна в пространствах $L^2(\mathcal{X})$ и $C(\mathcal{X})$ и $C_1 = \sqrt{2\left(1 + \|k\|^2_{C(\mathcal{X}^2)} \|\mathbf{K}^{-1}_{mm}\|^2_{L(\ell^\infty;\ell^1)}\right)}$.

Доказательство. В силу неравенства треугольника имеем

$$W_{2,L^2(\mathcal{X})}\left(f^{(d)}, f \mid \boldsymbol{y}\right) \leq W_{2,L^2(\mathcal{X})}\left(f^{(d)}, f^{(s)}\right) + W_{2,L^2(\mathcal{X})}\left(f^{(s)}, f \mid \boldsymbol{y}\right)$$

Далее мы последовательно раскрываем первое слагаемое поэлементно. Для произвольного x имеем

$$\begin{aligned}
\left|f^{(d)}(x) - f^{(s)}(x)\right|^2 &\leq 2\left(\left|f^{(w)}(x) - f(x)\right|^2 + \left|\mathbf{K}_{xm}\mathbf{K}^{-1}_{mm}\left(f^{(w)}(\boldsymbol{z}) - f(\boldsymbol{z})\right)\right|^2\right) \\
&\leq 2\left(\left\|f^{(w)} - f\right\|^2_{L^\infty(\mathcal{X})} + \left\|\mathbf{K}_{xm}\mathbf{K}^{-1}_{mm}\right\|^2_{\ell^1} \left\|f^{(w)}(\boldsymbol{z}) - f(\boldsymbol{z})\right\|^2_{\ell^\infty}\right) \\
&\leq 2\left(\left\|f^{(w)} - f\right\|^2_{L^\infty(\mathcal{X})} + \|\mathbf{K}_{xm}\|^2_{\ell^\infty} \left\|\mathbf{K}^{-1}_{mm}\right\|^2_{L(\ell^\infty;\ell^1)} \left\|f^{(w)} - f\right\|^2_{L^\infty}(\mathcal{X})\right) \\
&\leq 2\left(1 + \|k\|^2_{C(\mathcal{X}^2)} \left\|\mathbf{K}^{-1}_{mm}\right\|^2_{L(\ell^\infty;\ell^1)}\right) \left\|f^{(w)} - f\right\|^2_{L^\infty(\mathcal{X})} \\
&= 2\left(1 + \|k\|^2_{C(\mathcal{X}^2)} \left\|\mathbf{K}^{-1}_{mm}\right\|^2_{L(\ell^\infty;\ell^1)}\right) \left\|f^{(w)} - f\right\|^2_{C(\mathcal{X})}
\end{aligned}$$

В первой строчке мы воспользовались правилом Маферона, во второй неравенством Гёльдера для $p = 1, q = \infty$, в третьей использовали определение операторной нормы. Чтобы получить искомую оценку, проинтегрируем получившееся выражение:

$$\begin{aligned}
W^2_{2,L^2(\mathcal{X})}\left(f^{(d)}, f^{(s)}\right) &= \inf_{\gamma \in \mathcal{C}} \int \left\|f^{(d)} - f^{(s)}\right\|^2_{L^2(\mathcal{X})} \, \mathrm{d}\gamma \\
&\leq C | \inf_{\gamma \in \mathcal{C}} \int \left\|f^{(w)} - f\right\|^2_{C(\mathcal{X})} \, \mathrm{d}\gamma \\
&= C W^2_{2,C(\mathcal{X})}\left(f^{(w)}, f\right),
\end{aligned}$$

где $C = 2\left(1 + \|k\|^2_{C(\mathcal{X}^2)} \|\mathbf{K}^{-1}_{mm}\|^2_{L(\ell^\infty;\ell^1)}\right)$. Тогда, снимая квадрат, получим требуемое выражение. $\square$