arXiv:2602.13871v1 [math.ST] 14 Feb 2026

集成条件高斯过程（Ens-CGP）：表示、几何与推断

Sai Ravela Jae Deok Kim Kenneth Gee Xingjian Yan Samson Mercier
LubnaAlbarghouty Anamitra Saha

地球信号与系统组 地球、大气与行星
科学系 麻省理工学院

摘要

我们提出集成条件高斯过程（Ens-CGP），这是一种有限维综合框架，以条件高斯分布为集合推断的核心。条件高斯过程（CGP）直接由高斯过程经条件化得到；在线性高斯情形下，它为高斯先验和线性观测定义了完整的后验分布。经典卡尔曼滤波是一种递归算法，在动力学假设下计算同一条件分布；因此，条件高斯分布本身是底层的表示对象，而滤波器则是其一种计算实现。从这个意义上说，CGP 为卡尔曼类方法以及等价形式提供了概率基础，这些等价形式包括严格凸二次规划（最大后验估计）、再生核希尔伯特空间正则化回归和经典正则化。

Ens-CGP 是该对象的集成实例化：通过将经验集合矩视为一个（可能是低秩的）高斯先验，并执行精确条件化而获得。通过将表示（GP → CGP → Ens-CGP）与计算（卡尔曼滤波器、集合卡尔曼滤波变体和迭代集合方案）分离开来，该框架将一个较早确立的推断表示基础与集成导出的先验联系起来，并阐明了概率、变分和集成视角之间的关系。

1 引言

我们介绍集成条件高斯过程（Ens-CGP），这是一种用于高斯推断的有限维框架，其中集合统计量定义先验分布，而条件化产生后验分布，从而统一了高斯过程、再生核希尔伯特空间回归、二次规划和卡尔曼类更新。

有一个数学对象贯穿于几种文献领域，而这些领域在教学和实践中往往被视为彼此不同：高斯条件化（高斯过程与克里金法） [29], 核/再生核希尔伯特空间正则化（样条与学习理论） [40], 二次优化（最大后验估计量与

Tikhonov 型程序）[37], 以及*Kalman* 式更新（数据同化中的卡尔曼滤波/集合卡尔曼滤波）[12, 20]。在线性高斯设定下，这些并非仅仅彼此相关，而是同一个推断问题的不同表述。然而，在大规模实践中（尤其是使用集合时），这种等价关系往往没有被明确说明，而算法层面的叙述（例如集合卡尔曼滤波的变体、迭代集合方案）可能会模糊究竟表示了什么、又近似了什么。

核心主张（等价关系图）。固定有限离散化和带高斯噪声的线性观测模型。状态向量的高斯先验会诱导状态与观测的联合高斯分布，条件化后得到通过舒尔补公式 [4, 37] 求得的后验高斯分布。这个后验分布在离散形式下称为条件高斯过程（CGP），在假设线性高斯的情况下，它与卡尔曼分析后验完全一致：后验均值和后验协方差均相同。

后验均值同时也是：(i) 卡尔曼滤波器／平滑器的均值更新；(ii) 严格凸二次规划（负对数后验）的唯一极小化解；以及 (iii) 再生核希尔伯特空间正则化回归问题的解，该问题的范数由同一个协方差／核矩阵诱导。

后验协方差等于可容许子空间上受限的海森矩阵／精度矩阵的逆（若矩阵奇异，则为 Moore–Penrose 伪逆），由此将概率条件化与优化几何联系起来 [9]。此外，协方差的任何平方根分解——无论是谱模态、特征映射还是集合异常矩阵——都会诱导相同的子空间几何，因此得到相同的后验分布 [39]。

本文明确阐述了这种等价关系在有限维向量空间中的成立。

GP/CGP 框架为何重要。 GP/CGP 视角以表示为先：它规定了一个联合高斯分布，因此也确定了条件化后的整个后验分布。这不仅确定了一个估计量，还确定了其不确定性以及更新发生的结构。特别是，对联合高斯分布进行条件化会同时得到后验均值和后验协方差 [29]；卡尔曼滤波器则是在这一线性高斯框架 [32]下对同一条件化恒等式的递归实现。因此，条件高斯过程并非卡尔曼滤波的替代方案，而是卡尔曼滤波所计算的概率对象。

然而，与经典滤波器表述不同，条件高斯过程并不依赖于动力学模型或序贯数据同化。对于任意选定的高斯先验和线性信息算子（等价地，对于任意联合高斯模型，包括高斯图模型），它都是一次条件化操作。一旦将条件分布视为基本对象，集合构造、局地化、膨胀和迭代反演方案便可理解为对该分布的逼近或修改 [10]。将表示（条件高斯分布）与算法（计算方法）区分开来，对于清晰阐释、归因和解释至关重要。

条件高斯过程适合作为自然的首要对象。如果首要目标是概率推断、条件化、不确定性传播、将信息归因于数据而非先验，或与下游决策相结合，那么条件高斯过程/Ens-CGP 就是自然的顶层对象。如果首要目标是确定性正则化理论、逼近类、收敛速率、表示定理和极小极大分析，那么再生核希尔伯特空间视角可能是更方便的出发点。等价关系图表明，在线性高斯情形下，这些核心视角彼此一致；选择取决于侧重点，而非数学本身。

历史视角。 空间场的高斯条件化在地质统计学中表现为克里金 [25]，在概率论中则表现为高斯随机场 [2]。再生核希尔伯特空间–贝叶斯等价性体现在样条理论和正则化（Kimeldorf–Wahba/Wahba） [22]中，也体现在由核构造再生核希尔伯特空间的 Moore–Aronszajn 构造 [6]中。Kalman 的贡献在于递归地利用线性动力学模型 [20]中的相同条件化恒等式。在现代机器学习中，“高斯过程回归”普及了“核即先验”的叙事 [28]。我们的目标不是重新命名这些结果，而是以一种可直接用于大规模离散化计算和集合实现的形式，揭示它们共同的基本结构。

贡献与结构。 本文围绕一条主线组织全文：仅在必要时从经典基础出发，最终落到一个完全离散的向量空间：

- 我们陈述离散等价图 （条件高斯过程 $\Leftrightarrow$二次规划（最大后验估计） $\Leftrightarrow$ 再生核希尔伯特空间回归 $\Leftrightarrow$ 卡尔曼滤波/集合卡尔曼滤波均值更新），并在适当处理奇异协方差和满足值域约束的情况下证明该等价性。
- 我们通过平方根和格拉姆矩阵，将核／协方差与几何联系起来。
- 我们定义了*Ens-CGP*：通过条件化一个高斯先验得到的条件高斯分布，其中均值和协方差由集合根据经验定义。这将集合表示的内容（低秩先验）与如何计算（集合卡尔曼滤波变体及其他方案）区分开来。
- 我们从表示层面解释，为什么在没有新信息的情况下反复使用同一观测，不能被解释为卡尔曼条件化，除非明确指定引入新信息的概率模型；这也阐明了迭代集合反演的论述在哪些方面偏离了贝叶斯更新。

高斯条件化、舒尔补恒等式、卡尔曼滤波器，以及贝叶斯线性反问题、二次规划与再生核希尔伯特空间正则化之间的等价关系，都是经典结果 [32, 40]。这些数学恒等式都不是新的。新颖之处在于将集合明确纳入条件高斯框架。我们将集合–条件高斯过程 ***(Ens-CGP)*** 引入为一种表示对象：其先验均值和协方差由一个集合通过经验方式定义的条件高斯分布。在这一表述中，

集合卡尔曼方法并非基本构造，而是底层条件高斯分布的计算实现 [39]。集合成为一种低秩先验表示，而条件化则产生定义良好的后验分布。

这一重新表述将 EnKF 类型的方法纳入条件高斯过程，而不是将条件高斯过程视为集合卡尔曼滤波的一种特殊情况。鉴于其关于历史优先权的主张，条件高斯过程成为一种基础性的概率论与推断理论对象；集合则被理解为其有限维实例，而卡尔曼相关方法则是特殊情况。因此，本文的贡献在于结构层面：厘清层级关系，将表示与算法区分开来，并明确了二者的等价关系

$$\text{条件高斯过程} \equiv \text{最大后验估计（二次规划）} \equiv \text{再生核希尔伯特空间正则化} \equiv \text{卡尔曼后验（均值和协方差）},$$

Ens-CGP 是同一基础框架以集合形式实例化后的形态。

Ens-CGP 为一系列目前正在开发的逆问题与学习问题提供了共同的表示框架，包括方程发现 [14], 地震震相关联 [3], 大地电磁反演 [21], 气候降尺度 [31], 代理建模 [26], 以及天体物理推断。

本文其余部分将介绍论证离散化所需的最简经典背景知识（黎斯、协方差算子、Mercer/Karhunen–Loève、Moore–Aronszajn），随后回到支撑计算与集合方法的有限维框架。

记号与假设

概率。 所有随机变量都定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上。期望和协方差分别记作 $\mathbb{E}[\cdot]$ 和 $\mathrm{Cov}(\cdot,\cdot)$。

指标集与采样。 设 $\mathcal{X}$ 为一个指标集（例如空间位置），并令 $f = \{f(x) : x \in \mathcal{X}\}$ 表示一个随机过程（这是一种表示工具；计算仅在固定的有限离散化 X上进行）。对于有限子集 $X = \{x_1, \ldots, x_n\} \subset \mathcal{X}$，我们记作

$$\boldsymbol{f}_X := (f(x_1), \ldots, f(x_n))^\top \in \mathbb{R}^n, \qquad \boldsymbol{m}_X := \mathbb{E}[\boldsymbol{f}_X], \qquad (\boldsymbol{K}_{XX})_{ij} := \mathrm{Cov}(f(x_i), f(x_j)).$$

当子集 t X 根据上下文固定后，我们简写为 $\boldsymbol{f} := \boldsymbol{f}_X$, $\boldsymbol{m} := \boldsymbol{m}_X$，以及 $\boldsymbol{K} := \boldsymbol{K}_{XX}$.

高斯过程作为有限维高斯向量的一致族的形式化定义见第3节。下文中，“条件高斯过程”表示对这样的有限维联合高斯向量进行条件化。

线性观测模型（离散）。我们考虑

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}), \qquad \boldsymbol{\varepsilon} \perp \boldsymbol{f}, \tag{1}$$

其中 $\boldsymbol{H} \in \mathbb{R}^{m\times n}$ 和 $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ 均为对称正定（SPD），除非另有说明。

半正定与伪逆。 $\boldsymbol{K} \succeq 0$ 表示对称半正定（PSD）。 $\boldsymbol{K}^{\dagger}$ 是 Moore–Penrose 伪逆。

加权范数。 对于对称正定的 $\boldsymbol{A}$，我们记 $\|\boldsymbol{v}\|_{\boldsymbol{A}}^2 := \boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v}$。

2 预览：离散等价路线图

我们的目标是表明，若干标准对象在不同表示下其实是同一个 推断问题。本文的讨论是有限维且注重计算的：我们使用向量、矩阵以及（可能是低秩的）协方差。本文后续的经典理论章节将阐释这些对象的动机，而下文的等价关系则构成实际应用的最终形式。

阅读指南。 我们先给出完全有限维的等价关系图（本节），因为它是计算上的最终形式。接下来的经典理论章节将说明这些矩阵如何从连续对象中产生，随后我们将回到针对集合的纯离散构造。

定义 1 (条件高斯过程（有限维，条件高斯分布）**)**。 设 $\boldsymbol{f} \in \mathbb{R}^n$ 和 $\boldsymbol{y} \in \mathbb{R}^m$ 联合服从高斯分布。那么 $\boldsymbol{f} \mid \boldsymbol{y}$ 也服从高斯分布，其均值和协方差由 *Schur* 补条件化公式给出。本文中，*"CGP"*指固定离散化 X 上的这一条件高斯分布；高斯过程（*GP*）的过程级定义将在第*3*节回顾。

定理 1 （CGP $=$经典二次型（最大后验估计））。设 $\boldsymbol{f} \in \mathbb{R}^n$ 的先验为 $\boldsymbol{f} \sim \mathcal{N}$（$\boldsymbol{m}$、$\boldsymbol{K}$），其中 $\boldsymbol{K} \succeq 0$，观测 $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\varepsilon}$ 满足 $\boldsymbol{\varepsilon} \sim \mathcal{N}$（$\boldsymbol{0}$、$\boldsymbol{R}$），$\boldsymbol{R} \succ 0$，且与 $\boldsymbol{f}$独立。定义位移变量

$$\boldsymbol{x} := \boldsymbol{f} - \boldsymbol{m} \in \mathbb{R}^n, \qquad \boldsymbol{d} := \boldsymbol{y} - \boldsymbol{H}\boldsymbol{m} \in \mathbb{R}^m.$$

然后，负对数后验（忽略加性常数）为凸二次函数

$$\mathcal{J}(\boldsymbol{x}) = \frac{1}{2}\,\boldsymbol{x}^\top \boldsymbol{Q}\,\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + c, \qquad \boldsymbol{x} \in \mathrm{Range}(\boldsymbol{K}), \tag{2}$$

with

$$\boldsymbol{Q} := \boldsymbol{K}^{\dagger} + \boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H} \in \mathbb{R}^{n\times n}, \tag{3}$$

$$\boldsymbol{q} := -\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{d} \in \mathbb{R}^n, \tag{4}$$

$$c := \frac{1}{2}\,\boldsymbol{d}^\top \boldsymbol{R}^{-1}\boldsymbol{d} \in \mathbb{R}. \tag{5}$$

如果 $\boldsymbol{K} \succ 0$，那么 $\boldsymbol{x} \in \mathbb{R}^n$ 和 $\boldsymbol{Q} \succ 0$；如果 $\boldsymbol{K} \succeq 0$ 是奇异的，则目标函数仅在值域($\boldsymbol{K}$).

(2) 的唯一极小化解 $\boldsymbol{x}_\star$ 满足正规方程

$$\boldsymbol{Q}\boldsymbol{x}_\star = -\boldsymbol{q}, \tag{6}$$

且后验均值为 $\boldsymbol{m}_{\text{post}} = \boldsymbol{m} + \boldsymbol{x}_\star$ *[38]*。此外， $\boldsymbol{m}_{\text{post}}$ 等于条件均值$\mathbb{E}[\boldsymbol{f} \mid \boldsymbol{y}]$。

定理 2 （二次规划 = 再生核希尔伯特空间正则化回归（Wahba，离散情形））。 设 $\boldsymbol{K} \succeq 0$并定义有限维希尔伯特空间

$$\mathcal{H}_{\boldsymbol{K}} := \text{Range}(\boldsymbol{K}) \subseteq \mathbb{R}^n, \qquad \langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{H}_{\boldsymbol{K}}} := \boldsymbol{u}^\top \boldsymbol{K}^\dagger \boldsymbol{v}.$$

那么，定理 *1* 中的极小值点等价地为

$$\boldsymbol{m}_{\text{post}} = \underset{\boldsymbol{g} \in \boldsymbol{m} + \mathcal{H}_{\boldsymbol{K}}}{\arg\min} \ \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{g}\|^2_{\boldsymbol{R}^{-1}} + \|\boldsymbol{g} - \boldsymbol{m}\|^2_{\mathcal{H}_{\boldsymbol{K}}},$$

即，条件高斯过程推断 等于由 $\boldsymbol{K}$ *[40]*诱导的再生核希尔伯特空间范数下的再生核希尔伯特空间正则化回归。

定理 3 （平方根、格拉姆矩阵与再生核希尔伯特空间几何（离散情形））。 设 $\boldsymbol{K} \in \mathbb{R}^{n \times n}$ 为半正定矩阵，秩为 r。那么存在一个特征／平方根因子 $\boldsymbol{A} \in \mathbb{R}^{n \times r}$，使得

$$\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}^\top.$$

*De*定义特征矩阵 $\boldsymbol{\phi} := \boldsymbol{A}^\top \in \mathbb{R}^{r \times n}$，并令 $\boldsymbol{\phi}_i$ 表示其第 i 列。则

$$\boldsymbol{K} = \boldsymbol{\phi}^\top \boldsymbol{\phi}, \qquad K_{ij} = \boldsymbol{\phi}_i^\top \boldsymbol{\phi}_j,$$

因此 $\boldsymbol{K}$ 是特征向量 $\{\boldsymbol{\phi}_i\}_{i=1}^n$ *[34, 35]* 的格拉姆矩阵 。对偶（特征空间）格拉姆矩阵为 $\boldsymbol{G}$: $= \boldsymbol{\phi}\boldsymbol{\phi}^\top = \boldsymbol{A}^\top \boldsymbol{A}$。此外，诱导出的有限维再生核希尔伯特空间为 *[7]*

$$\mathcal{H}_{\boldsymbol{K}} = \text{Range}(\boldsymbol{K}) = \text{Range}(\boldsymbol{A}), \qquad \langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{H}_{\boldsymbol{K}}} = \boldsymbol{u}^\top \boldsymbol{K}^\dagger \boldsymbol{v}.$$

推论 1 （线性高斯假设下 Ens-CGP 与集合卡尔曼滤波的均值一致）。 如果 （$\boldsymbol{f}$、 $\boldsymbol{K}$）由一个集合估计，并且更新在卡尔曼增益中使用 $\boldsymbol{K}$

$$\boldsymbol{G} := \boldsymbol{K}\boldsymbol{H}^\top (\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R})^{-1},$$

那么，后验均值更新 $\boldsymbol{f} \mapsto \boldsymbol{f} + \boldsymbol{G}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f})$ 正是第 *11* 节所定义的集合条件高斯过程的条件均值。标准集合卡尔曼滤波更新是蒙特卡罗实现，其样本均值等于此更新（但存在抽样误差，以及局地化／膨胀等其他数值选择带来的影响） *[12]*。

3 高斯过程：经典定义

高斯过程最好被理解为一族相容的 有限维高斯向量。这既是其形式定义，也是我们采取离散化优先立场的依据。

3.1 定义（作为一致的有限维高斯分布的过程）

一个高斯过程（*GP*）是由实随机变量 $\{f(x) : x \in \mathcal{X}\}$ 构成的集合，使得对于任意有限子集 $X = \{x_1, \dots, x_n\}$，向量 $\boldsymbol{f}_X$ 服从多元正态分布：

$$\boldsymbol{f}_X \sim \mathcal{N}(\boldsymbol{m}_X, \boldsymbol{K}_{XX}), \quad (\boldsymbol{m}_X)_i = \mathbb{E}[f(x_i)], \quad (\boldsymbol{K}_{XX})_{ij} = \mathrm{Cov}(f(x_i), f(x_j)).$$

3.2 一致性与柯尔莫哥洛夫延拓

若一个族 $\{\mathcal{N}(\boldsymbol{m}_X, \boldsymbol{K}_{XX})\}_X$ 在边缘化下保持一致，则它定义了一个随机过程：对于 $Y \subset X$，将 $\mathcal{N}(\boldsymbol{m}_X, \boldsymbol{K}_{XX})$ 边缘化到由 Y 索引的坐标上，所得边缘分布等于 $\mathcal{N}(\boldsymbol{m}_Y, \boldsymbol{K}_{YY})$。对于高斯族，这一性质自动成立，因为多元高斯分布的边缘分布仍为高斯分布，且对应于相应的子向量和子矩阵 [29]。

定理4 （柯尔莫哥洛夫延拓定理）。 设 $\{P_X\}_{X \subset \mathcal{X},\ |X| < \infty}$ 是定义在 $\mathbb{R}^{|X|}$ 上的一族概率测度，且在边缘化下保持一致。那么，存在一个随机过程 $\{f(x) : x \in \mathcal{X}\}$，定义在某个 $(\Omega, \mathcal{F}, \mathbb{P})$ 上，使得对于每个有限 $X = \{x_1, \dots, x_n\}$ *[8]*，

$$(f(x_1), \dots, f(x_n)) \sim P_X.$$

注 1 （为何强调离散情形）。 柯尔莫哥洛夫延拓保证过程存在，但不保证样本路径正则性（连续性、可微性等）。由于本文中的所有计算都在固定的有限离散化 X 上进行，我们将高斯过程视为有限维高斯向量族，并将无限维过程看作一种表示工具，而非计算对象。在全文中，我们最终都将范围限制在固定有限集合 X 上，因此高斯过程完全由随机向量 $\boldsymbol{f} \in \mathbb{R}^n$ 和参数 $(\boldsymbol{m}, \boldsymbol{K})$ 表示。

4 线性泛函与里斯表示

本节阐述连续线性测量如何导出内积，以及这为何对协方差至关重要。其关键结果是一个算子恒等式，离散化后得到矩阵协方差[36]。

4.1 希尔伯特空间与线性泛函

设 $\mathcal{H}$ 是一个实希尔伯特空间，$\ell : \mathcal{H} \to \mathbb{R}$ 是一个有界线性泛函。

定理 5（里斯表示）。 存在唯一的 $h_\ell \in \mathcal{H}$，使得 $\ell(g) = \langle g,\ h_\ell \rangle_{\mathcal{H}}$ 对所有$g \in \mathcal{H}$ *[30]*成立。

4.2 协方差算子

设 f 是一个均值为 $m = \mathbb{E}[f]$（博赫纳积分）的 $\mathcal{H}$ 值随机元。定义秩一（外积）算子

$$(u \otimes v)\, g := u\, \langle v, g \rangle_{\mathcal{H}}, \qquad u, v, g \in \mathcal{H}.$$

协方差算子 $\mathcal{C} : \mathcal{H} \to \mathcal{H}$ 为

$$\mathcal{C} := \mathbb{E}\big[(f - m) \otimes (f - m)\big],$$

等价地，可由下式刻画

$$\langle h, \mathcal{C} g \rangle_{\mathcal{H}} = \mathbb{E}\Big[\, \langle h, f - m \rangle_{\mathcal{H}}\ \langle f - m, g \rangle_{\mathcal{H}} \,\Big], \qquad \forall\, h, g \in \mathcal{H}.$$

那么，对于有界线性泛函 $\ell_1,\ \ell_2$（其表示元为 $h_{\ell_1},\ h_{\ell_2}$），

$$\mathrm{Cov}(\ell_1(f), \ell_2(f)) = \langle h_{\ell_1}, \mathcal{C} h_{\ell_2} \rangle_{\mathcal{H}}.$$

注 2. 这个算子恒等式是矩阵协方差 *[1]* 的连续类比。它是进行对角化（*Mercer/Karhunen–Loève*）和离散化（第 *8* 节）的正确对象。

5 默瑟理论与 Karhunen–Lo`eve 展开

我们将核/协方差与谱分解联系起来。在离散情形下，这转化为特征值/SVD 分解，这也是集合平方根 [39]的基础。

5.1 积分算子

假设 $\mathcal{X}$ 是紧的，带有测度 μ，并且 k 是连续、对称、正定的。定义T_k： $L^2(\mathcal{X},\ \mu) \to L^2(\mathcal{X},\ \mu)$，如下所示

$$(T_k g)(x) = \int_{\mathcal{X}} k(x, x')\, g(x')\, d\mu(x').$$

那么 T_k 是紧、自伴且正的。

5.2 默瑟展开

假设 k 在配备有限 Borel测度 μ 的紧空间 $\mathcal{X}$ 上连续、对称且正定。那么 Mercer定理给出特征对 $(\lambda_i,\ \psi_i)$，其中 $\lambda_i \geq 0$ 且 $\{\psi_i\}$

在 $L^2(\mathcal{X},\ \mu)$ 中正交归一，使得

$$k(x,x') = \sum_{i=1}^{\infty} \lambda_i\, \psi_i(x)\psi_i(x'),$$

其中该级数在 $\mathcal{X}\times\mathcal{X}$ [1, 7] 上绝对且一致收敛。

5.3 Karhunen–Lo`eve（KL）表示

如果 f 是定义在 $\mathcal{X}$ 上、协方差为 k 的中心化高斯过程，并且满足温和的可积性条件

$$\int_{\mathcal{X}} k(x,x)\, d\mu(x) < \infty$$

则有

$$f(x) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, z_i\, \psi_i(x), \qquad z_i \overset{iid}{\sim} \mathcal{N}(0,1),$$

其中（$\lambda_i,\ \psi_i$）是协方差（积分）算子 T_k 在 L^2（$\mathcal{X},\ \mu$）上的特征对。对于每个固定的 $x\in\mathcal{X}$，该级数在 $L^2(\Omega)$ 中收敛（在更强的条件下，还可以按逐点或一致意义收敛） [1, 15]。

6 Moore–Aronszajn 与再生核希尔伯特空间

本节解释核为何定义希尔伯特空间几何和再生性质 [6]。在离散情形下，再生核希尔伯特空间几何正是由 $\boldsymbol{K}^{\dagger}$ 在值域（$\boldsymbol{K}$）上诱导的几何 [7]。

6.1 正定核

一个对称函数 ion k： $\mathcal{X}\times\mathcal{X}\to\mathbb{R}$ 若对于所有有限的 $\{x_i\}$ 及系数 $\{a_i\}$，则称其为正定的 ，

$$\sum_{i,j} a_i a_j k(x_i,x_j) \geq 0.$$

6.2 Moore–Aronszajn

定理 6 （Moore–Aronszajn）。 对于每个正定核 k，都存在唯一的再生核希尔伯特空间$\mathcal{H}_k$，其为定义在 $\mathcal{X}$上的函数空间，并满足 k（$\cdot,\ x$）$\in\mathcal{H}_k$，且再生性质成立 *[6]*:

$$f(x) = \langle f, k(\cdot,x)\rangle_{\mathcal{H}_k} \quad \forall f\in\mathcal{H}_k.$$

6.3 特征映射和格拉姆矩阵

定义典范特征映射 $\phi(x) := k(\cdot,x)\in\mathcal{H}_k$。于是

$$k(x,x') = \left\langle \phi(x), \phi(x')\right\rangle_{\mathcal{H}_k}.$$

对于有限集合 $X = \{x_1, \dots, x_n\}$，格拉姆矩阵满足

$$(\boldsymbol{K}_{XX})_{ij} = k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_{\mathcal{H}_k}.$$

在完全离散向量情形下，我们可以选取一个平方根 $\boldsymbol{A}$，使得 $\boldsymbol{K}_{XX} = \boldsymbol{A}\boldsymbol{A}^\top$，并定义$\boldsymbol{\phi} := \boldsymbol{A}^\top$，从而 $\boldsymbol{K}_{XX} = \boldsymbol{\phi}^\top \boldsymbol{\phi}$（定理 3） [34, 7]。

7 有限维中的 Wahba等价性

现在，我们明确说明离散情形下的核心推断等价性。这是概率（条件化）、优化（最大后验估计）与几何（再生核希尔伯特空间）之间的实践桥梁 [27, 40]。

7.1 先验与似然（线性高斯）

设 $\boldsymbol{f} \in \mathbb{R}^n$，并假设高斯先验

$$\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{K}), \qquad \boldsymbol{K} \succeq 0,$$

以及观测模型（1）。那么联合向量

$$\begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{y} \end{bmatrix}$$

服从高斯分布，且

$$\mathbb{E}[\boldsymbol{y}] = \boldsymbol{H}\boldsymbol{m}, \qquad \mathrm{Cov}(\boldsymbol{y}, \boldsymbol{y}) = \boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R}, \qquad \mathrm{Cov}(\boldsymbol{f}, \boldsymbol{y}) = \boldsymbol{K}\boldsymbol{H}^\top.$$

7.2 条件高斯过程：向量的条件高斯分布

本节针对线性高斯模型 (1) 实例化定义 1。标准高斯条件化可得 [29, 32]

$$\boldsymbol{m}_{\mathrm{post}} = \boldsymbol{m} + \boldsymbol{K}\boldsymbol{H}^\top (\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R})^{-1} (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{m}) \tag{7}$$

and

$$\boldsymbol{K}_{\mathrm{post}} = \boldsymbol{K} - \boldsymbol{K}\boldsymbol{H}^\top (\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R})^{-1} \boldsymbol{H}\boldsymbol{K}. \tag{8}$$

后验协方差作为逆海森矩阵（离散条件高斯过程）。在线性高斯框架下，先验为 $\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{m},\ \boldsymbol{K})$（可能是奇异的），观测模型为 $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0},\ \boldsymbol{R})$，负对数后验（相差一个加法常数）是二次型

$$\mathcal{J}(\boldsymbol{f}) = \frac{1}{2}\|\boldsymbol{f} - \boldsymbol{m}\|^2_{\boldsymbol{K}^\dagger} + \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f}\|^2_{\boldsymbol{R}^{-1}}, \qquad \boldsymbol{f} \in \boldsymbol{m} + \mathrm{Range}(\boldsymbol{K}).$$

其海森矩阵（可容许子空间上的精度矩阵）为

$$\boldsymbol{Q} = \boldsymbol{K}^{\dagger} + \boldsymbol{H}^{\top}\boldsymbol{R}^{-1}\boldsymbol{H}.$$

后验协方差等于 $\boldsymbol{Q}$ 在 Range($\boldsymbol{K}$) 上的逆矩阵；等价地， $\boldsymbol{K}_{\mathrm{post}}$ 是 $\boldsymbol{Q}$ 在 Range($\boldsymbol{K}$) 上的 Moore–Penrose 伪逆。这与对联合高斯分布进行条件化得到的 Schur 补协方差一致。

7.3 证明条件高斯过程 = 二次规划（最大后验估计） = 再生核希尔伯特空间（有限维）

证明。 令 $\boldsymbol{f} = \boldsymbol{m} + \boldsymbol{x}$ 和 $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{m}$ 满足 $\boldsymbol{y} - \boldsymbol{H}\boldsymbol{f} = \boldsymbol{d} - \boldsymbol{H}\boldsymbol{x}$。忽略与 $\boldsymbol{x}$ 无关的加法常数，负对数后验等于

$$\mathcal{J}(\boldsymbol{x}) = \frac{1}{2}\,\|\boldsymbol{d} - \boldsymbol{H}\boldsymbol{x}\|^2_{\boldsymbol{R}^{-1}} + \frac{1}{2}\,\|\boldsymbol{x}\|^2_{\boldsymbol{K}^{\dagger}}, \qquad \boldsymbol{x} \in \mathrm{Range}(\boldsymbol{K}),$$

当 $\boldsymbol{K}$ 为奇异矩阵时，必须限制定义域。

展开后得到

$$\mathcal{J}(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\top}(\boldsymbol{K}^{\dagger} + \boldsymbol{H}^{\top}\boldsymbol{R}^{-1}\boldsymbol{H})\boldsymbol{x} + (-\boldsymbol{H}^{\top}\boldsymbol{R}^{-1}\boldsymbol{d})^{\top}\boldsymbol{x} + \frac{1}{2}\boldsymbol{d}^{\top}\boldsymbol{R}^{-1}\boldsymbol{d},$$

这正是代入 (3)–(5) 后的 (2)。在 Range($\boldsymbol{K}$) 上的一阶最优性条件为 $\boldsymbol{Q}\boldsymbol{x}_{\star} = -\boldsymbol{q}$。

由于模型是线性高斯的，后验为高斯分布，其均值等于最大后验估计（负对数后验的极小化解），因此 $\boldsymbol{m}_{\mathrm{post}} = \boldsymbol{m} + \boldsymbol{x}_{\star} = \mathbb{E}[\boldsymbol{f} \mid \boldsymbol{y}]$。最后，定理 2 中的再生核希尔伯特空间正则化形式由恒等式 $\|\boldsymbol{g} - \boldsymbol{m}\|^2_{\mathcal{H}_{\boldsymbol{K}}} = \|\boldsymbol{g} - \boldsymbol{m}\|^2_{\boldsymbol{K}^{\dagger}}$ 在$\boldsymbol{m}$ + 值域($\boldsymbol{K}$)上成立（并且在其外部 $+\infty$） [7]。

□

8 离散化到向量空间

本节明确说明连续理论如何转化为矩阵，以及为何上述算子级恒等式会归约为有限维协方差。

点求值离散化。 给定 $X = \{x_1, \ldots, x_n\}$，定义 $\boldsymbol{f} = (f(x_1) \ldots, \ f(x_n))^{\top}$ 和$\boldsymbol{K}_{ij} = k(x_i, \ x_j)$。那么 $\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{m}, \ \boldsymbol{K})$ 恰好为 [29]。

一般线性泛函离散化。 设 ℓ_i 为有界线性泛函（例如，对传感器覆盖区域取平均）。定义

$$\boldsymbol{f}_{\ell} := (\ell_1(f), \ldots, \ell_n(f))^{\top}.$$

如果 f 是一个 G P（或高斯随机元）和 ℓ_i 有界，则 $\boldsymbol{f}_{\ell}$ 是高斯的，且

$$(\boldsymbol{K}_{\ell})_{ij} = \mathrm{Cov}(\ell_i(f), \ell_j(f)) = \left\langle h_{\ell_i}, \mathcal{C} h_{\ell_j} \right\rangle_{\mathcal{H}},$$

使用 Riesz 表示元和第 4 节中的协方差算子。这是离散化测量 [38]上的矩阵协方差。

9 平方根、模态与等价表示

所有核/协方差信息都可以通过平方根 [39]来表示。这是连接 KL/Mercer 模态、SVD/EOF 模态和集合异常矩阵 [19]的桥梁。

9.1 谱分解

对于半正定 $\boldsymbol{K}$，设 $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ 满足 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_r, 0, \dots)$，且 $\lambda_i > 0$ 适用于 $i \le r$。一种典范平方根为

$$\boldsymbol{K}^{1/2} := \boldsymbol{U}_r \boldsymbol{\Lambda}_r^{1/2} \in \mathbb{R}^{n \times r}, \qquad \boldsymbol{K} = \boldsymbol{K}^{1/2} (\boldsymbol{K}^{1/2})^\top.$$

9.2 任意平方根与奇异值分解

任何满足 $\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}^\top$ 的 $\boldsymbol{A} \in \mathbb{R}^{n \times r}$ 都是有效的平方根表示（例如，集合异常、Cholesky 因子、随机化草图）。其薄奇异值分解 $\boldsymbol{A} = \boldsymbol{U}_r \boldsymbol{\Sigma} \boldsymbol{V}^\top$ 意味着 $\boldsymbol{K} = \boldsymbol{U}_r \boldsymbol{\Sigma}^2 \boldsymbol{U}_r^\top$。因此，“集合平方根”和“EOF/奇异值分解模态”仅相差特征空间中的旋转 [23]。

定理 *3* 的证明。令 $\boldsymbol{K} \succeq 0$，并在其秩为 r 的值域上进行特征值分解 $\boldsymbol{K} = \boldsymbol{U}_r \boldsymbol{\Lambda}_r \boldsymbol{U}_r^\top$。令 $\boldsymbol{A} := \boldsymbol{U}_r \boldsymbol{\Lambda}_r^{1/2}$，使得 $\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}^\top$。定义 $\boldsymbol{\phi} := \boldsymbol{A}^\top$，得到 $\boldsymbol{K} = \boldsymbol{\phi}^\top \boldsymbol{\phi}$，其元素为 $K_{ij} = \boldsymbol{\phi}_i^\top \boldsymbol{\phi}_j$。最后，$\mathrm{Range}(\boldsymbol{K}) = \mathrm{Range}(\boldsymbol{A})$，且内积 $\boldsymbol{u}^\top \boldsymbol{K}^\dagger \boldsymbol{v}$ 在该值域上诱导出再生核希尔伯特空间几何。

□

10 集合近似

现在，我们用经验平均替代期望。此步骤将高斯过程/条件高斯过程转化为由集合实例化的先验/条件分布。

10.1 经验均值和协方差

给定一个集合 $\{\boldsymbol{f}^{(e)}\}_{e=1}^E \subset \mathbb{R}^n$，定义

$$\bar{\boldsymbol{f}} := \frac{1}{E} \sum_{e=1}^E \boldsymbol{f}^{(e)}, \qquad \boldsymbol{K} := \frac{1}{E-1} \sum_{e=1}^E (\boldsymbol{f}^{(e)} - \bar{\boldsymbol{f}})(\boldsymbol{f}^{(e)} - \bar{\boldsymbol{f}})^\top.$$

设异常矩阵为

$$\boldsymbol{A} := \frac{1}{\sqrt{E-1}} \big[\boldsymbol{f}^{(1)} - \bar{\boldsymbol{f}}, \dots, \boldsymbol{f}^{(E)} - \bar{\boldsymbol{f}} \big] \in \mathbb{R}^{n \times E}.$$

Then

$$\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}^\top, \qquad \mathrm{rank}(\boldsymbol{K}) \le E - 1, \qquad \mathrm{Range}(\boldsymbol{K}) = \mathrm{Range}(\boldsymbol{A}).$$

定义诱导特征矩阵

$$\boldsymbol{\phi} := \boldsymbol{A}^\top \in \mathbb{R}^{E\times n},$$

使得 $\boldsymbol{K} = \boldsymbol{\phi}^\top \boldsymbol{\phi}$ 是离散化索引 [17] 上的格拉姆矩阵。

这表明诱导离散再生核希尔伯特空间为 $\mathcal{H}_{\boldsymbol{K}}$ =Range($\boldsymbol{A}$)。

$\widehat{\ }$**经验协方差的记法**。我们用 $\boldsymbol{K}$ 表示经验协方差，并在强调其作为真实先验协方差估计值的作用时，将其与 $\boldsymbol{K}$ 互换使用；两者并无区别。

11 系综条件高斯过程（Ens-CGP）

现在，我们以一种纯离散且易于处理的方式定义核心对象：由集合定义的高斯先验及其条件高斯分布。

定义 2 (Ens-CGP)。固定一个有限离散化 $X = \{x_1, \dots, x_n\}$。设集合 $\{\boldsymbol{f}^{(e)}\}_{e=1}^{E} \subset \mathbb{R}^n$ 按第 *10* 节中的定义构成经验均值 $\boldsymbol{f}$ 和经验协方差 $\boldsymbol{K}$ 。在线性观测模型 (1) 下，系综条件高斯过程（Ens-CGP）是条件高斯分布

$$\boldsymbol{f} \mid \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{m}_{\text{post}}, \boldsymbol{K}_{\text{post}})$$

$\bar{\ }$通过将（7）–（8）应用于（$\boldsymbol{m}, \boldsymbol{K}$） = （$\boldsymbol{f}, \boldsymbol{K}$）以及增益

$$\boldsymbol{G} := \boldsymbol{K}\boldsymbol{H}^\top(\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R})^{-1}.$$

增益的适定性。即使先验协方差 $\boldsymbol{K}$ 秩亏（如集成近似中所示），矩阵

$$\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R}$$

也是对称正定的，因为 $\boldsymbol{R} \succ 0$。因此，逆矩阵存在，增益算子无需伪逆即可良好定义。

11.1 归因与几何

后验均值偏移满足

$$\boldsymbol{m}_{\text{post}} - \bar{\boldsymbol{f}} = \boldsymbol{K}\boldsymbol{H}^\top(\boldsymbol{H}\boldsymbol{K}\boldsymbol{H}^\top + \boldsymbol{R})^{-1}(\boldsymbol{y} - \boldsymbol{H}\bar{\boldsymbol{f}}) \in \text{Range}(\boldsymbol{K}) = \text{Range}(\boldsymbol{A}).$$

因此，后验调整仅限于集成张成空间；在 Ens-CGP 先验下，Null($\boldsymbol{K}$) 中的分量无法被检测到。

11.2 与卡尔曼滤波/EnKF 的联系（分析步）

在线性高斯模型下，条件化得到 (7)–(8)；这些恰好就是后验均值和协方差的 Kalman分析（观测更新）公式 [20, 32]。

在完整的卡尔曼滤波器中，预报 步会在分析步之前通过动力学模型传播 $(\boldsymbol{m}, \boldsymbol{K})$，并且新的观测会依次到达。Ens-CGP 对应于在给定先验 $(\boldsymbol{f}, \boldsymbol{K})$ 下的分析¯/条件化步；当 $\boldsymbol{K}$ 由集合估计时，所得均值更新与标准集合卡尔曼滤波均值更新一致（推论 1） [12]。额外的算法手段，例如局地化/膨胀，会修改 $\boldsymbol{K}$ ，从而修改所隐含的 Ens-CGP 先验；它们不属于定义 [5, 16]。

12 优先考虑高斯过程、条件高斯过程与 Ens-CGP，而非集合卡尔曼滤波和集合卡尔曼反演

本节从操作层面说明以表示为先的立场：条件高斯过程/Ens-CGP 定义了一个完整的概率对象（均值、协方差、几何以及二次目标函数），而许多迭代集合方案应被视为概率解释并不固定的优化启发式方法。

12.1 Kalman 作为 CGP 的特殊情形

卡尔曼滤波器通常被描述为一种用于线性动力系统 [4] 的状态估计算法。CGP 是高斯随机向量在线性信息条件下的条件分布；它同时给出 (i) 后验均值 和协方差，(ii) 精确的最大后验估计二次规划，以及 (iii) 其解析梯度和曲率。因此，卡尔曼滤波器虽然提供协方差更新，但 CGP 才是其背后的表示；当 CGP 的递推与动力学预测模型及序贯数据结合时，便得到卡尔曼滤波器。

定义负对数后验（相差一个加性常数）

$$\mathcal{J}(\boldsymbol{g}) := \tfrac{1}{2}\,\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{g}\|^2_{\boldsymbol{R}^{-1}} + \tfrac{1}{2}\,\|\boldsymbol{g} - \boldsymbol{m}\|^2_{\boldsymbol{K}^\dagger}, \qquad \boldsymbol{g} \in \boldsymbol{m} + \mathrm{Range}(\boldsymbol{K}). \tag{9}$$

那么，在 $\boldsymbol{m}$ +Range($\boldsymbol{K}$) 上的梯度和海森矩阵为

$$\nabla \mathcal{J}(\boldsymbol{g}) = -\boldsymbol{H}^\top \boldsymbol{R}^{-1}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{g}) + \boldsymbol{K}^\dagger(\boldsymbol{g} - \boldsymbol{m}), \tag{10}$$

$$\nabla^2 \mathcal{J}(\boldsymbol{g}) = \boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H} + \boldsymbol{K}^\dagger, \tag{11}$$

它在 $\boldsymbol{g}$ 上为常数（这是线性高斯情形的典型特征）。因此，CGP 不仅提供更新规则，还给出了其背后的完整二次曲面，以及后验协方差 (8)。

12.2 Ens-CGP：作为先验的集合

¯Ens-CGP 将 $(\boldsymbol{m},\ \boldsymbol{K})$ 替换为集合统计量 $(\boldsymbol{f},\ \boldsymbol{K})$。这使集合转化为具体的（通常是低秩的）高斯先验，而条件化则产生定义良好的后验。归因

陈述（发生了什么变化，以及变化发生在哪个子空间）成为关于诱导协方差几何及其值域的陈述。

12.3 集合　　在迭代反演中的作用：代理，而非 Kalman 性质

迭代集合反演方法（通常统称为“EnKI/EKI”）最好被理解为一种优化方案：它将经验协方差重复用作低秩度量/预条件器 [18]。

在这种作用下，集合主要用于近似数据失配目标函数的梯度和高斯–牛顿曲率，而不是从固定的贝叶斯后验 [33]中抽取样本。若没有明确的概率模型在每一步引入真正的新信息，这类迭代就不具备卡尔曼滤波性质：不存在一个被递归条件化的单一联合高斯分布，而且不断变化的集合协方差通常无法等同于固定似然下的后验协方差 [13]。

12.4　R重复使用固定观测并非贝叶斯条件化

对于反复重复使用同一观测 $\boldsymbol{y}$ 的迭代方案，一个核心的解释问题在于，贝叶斯条件化是一次单一更新： $\boldsymbol{f}|\boldsymbol{y}$。除非假设每次迭代都引入新的独立信息，否则反复使用 $\boldsymbol{y}$ 就等于重复计数 [24]。

命题 1（重复相同“观测”导致后验塌缩）。假设高斯先验 $\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{m}_0, \boldsymbol{K}_0)$ 具有 $\boldsymbol{K}_0 \succ 0$，且线性观测模型为 $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$ 具有$\boldsymbol{R} \succ 0$。如果（错误地）将同一个已实现的 $\boldsymbol{y}$ 视为 k 个相互独立且具有相同（$\boldsymbol{H}, \boldsymbol{R}$）的观测，那么所得后验协方差和均值等于

$$\boldsymbol{K}_k = \left(\boldsymbol{K}_0^{-1} + k\,\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{H}\right)^{-1}, \qquad \boldsymbol{m}_k = \boldsymbol{K}_k\left(\boldsymbol{K}_0^{-1}\boldsymbol{m}_0 + k\,\boldsymbol{H}^\top \boldsymbol{R}^{-1}\boldsymbol{y}\right). \tag{12}$$

特别　是，对于 $\boldsymbol{H} = \boldsymbol{I}$，有 $\boldsymbol{K}_k = (\boldsymbol{K}_0^{-1} + k\,\boldsymbol{R}^{-1})^{-1} \to \boldsymbol{0}$，且 $\boldsymbol{m}_k \to \boldsymbol{y}$ 作为 $k \to \infty$.

注释 3（对迭代集合方案的启示）。 式子(12) 并不是 单个观测下正确贝叶斯推断的性质。它是反复将同一数据实现视为新信息的结果。因此，任何反复重用 $\boldsymbol{y}$ 的算法，都必须被理解为一种正则化优化 （或是在经过明确修改的数据模型下进行推断）；这也立即解释了停止准则、阻尼及其他数值控制为何如此重要 *[33]*。这就是此类方案并非*“Kalman”*方法的含义，除非其迭代与一个定义明确、每一步都会引入新信息的概率模型相联系 *[13]*。从这一贝叶斯视角来看，此类迭代方案中的集合用于构造低秩几何（预条件／梯度近似），而非用来保证类似 *Kalman* 的后验传播。

13 历史注记：条件高斯分布与“条件高斯过程”

高斯向量和高斯场的条件化早于现代计算处理方法；这一研究对象由来已久，尽管“CGP”这一缩写并非如此。

以条件化为基本操作 若联合向量 $\begin{bmatrix} \boldsymbol{f}^\top & \boldsymbol{y}^\top \end{bmatrix}^\top$ 为高斯分布，则 $\boldsymbol{f} \mid \boldsymbol{y}$ 亦为高斯分布，其均值与协方差由舒尔补公式 (7)–(8) 给出。这是标准的多元正态理论。

从高斯随机场到克里金法 当 $\boldsymbol{f}$ 被解释为空间上的高斯随机场时，对含噪观测进行条件化即得到经典的克里金预测器 [11], ，它恰好是协方差由核指定的高斯模型的条件均值。

再生核希尔伯特空间与样条正则化 独立地，高斯条件化与惩罚最小二乘之间的等价性出现在平滑样条理论及其相关的再生核希尔伯特空间形式体系 [22, 40]中。用离散向量的语言来说，这正是定理 1 与定理 2。

现代机器学习术语 高斯过程回归推广了将核视为先验、将后验预测视为条件化的观点 [29]。在该文献中，“对高斯过程进行条件化”是标准操作；此处条件高斯过程指的是在离散化向量空间框架中由此得到的条件高斯分布。

14 首选构造：条件高斯过程还是再生核希尔伯特空间?

等价关系图表明，在线性高斯框架中，条件高斯过程（概率）与再生核希尔伯特空间正则化（几何）是同一个有限维对象的两个侧面。因此，两者的区别在于侧重点和可审计性。

14.1 两种等价的中心，各有不同侧重点

以再生核希尔伯特空间为先（以几何为先）。给定 $\boldsymbol{K} \succeq 0$，由此诱导的希尔伯特空间 $\mathcal{H}_{\boldsymbol{K}} = \mathrm{Range}(\boldsymbol{K})$及其内积 $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{H}_{\boldsymbol{K}}} = \boldsymbol{u}^\top \boldsymbol{K}^\dagger \boldsymbol{v}$ 将核/协方差转化为几何。正则化、表示子结构和特征分解用这种语言描述最为自然。

C以高斯过程为先（以推断为先）。条件高斯过程指定了整个条件分布

$$\boldsymbol{f} \mid \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{m}_{\mathrm{post}}, \boldsymbol{K}_{\mathrm{post}}).$$

这通过 $\boldsymbol{K}_{\mathrm{post}}$ 实现不确定性量化，并以规范形式定义二次最大后验目标函数 (9)，以及解析梯度 (10) 和曲率 (11)。因此，条件高斯过程将估计、不确定性、归因和优化统一于同一个对象中。

在基于集成的工作流中，可以明确命名并审计的对象是隐含的先验$^{-}$ $\widehat{}$分布 $\mathcal{N}(\boldsymbol{f},\ \boldsymbol{K})$，以及其在 $(\boldsymbol{H},\ \boldsymbol{R})$ 条件下的条件分布。

首选构造 Ens-CGP 是直接的表示基元：它唯一确定 (i) 均值更新、(ii) 后验协方差、(iii) 所诱导的再生核希尔伯特空间几何，以及 (iv) 最大后验估计优化问题。再生核希尔伯特空间仍然至关重要，但它是一种派生结构：它描述由集合协方差所诱导的几何。

然而，如果首要目标是确定性正则化理论（收敛率、逼近类、表示定理），而不确定性居于次要地位，那么再生核希尔伯特空间或许是更自然的顶层概念。当不确定性、条件化和归因处于首要地位时，条件高斯过程/Ens-CGP 则是更基础的框架。

15 结论

本文提炼出一条离散的表示主干：

$$\begin{aligned}&\text{条件高斯过程} \equiv \text{二次规划（最大后验估计）} \equiv \\ &\text{再生核希尔伯特空间正则化回归} \equiv \text{基于协方差} \\ &\text{核的条件化}\ \ \boldsymbol{K}.\end{aligned}$$

Ens-CGP 是同一对象在集合中的实例化：一个低秩高斯先验及其条件高斯分布。这将表示与算法区分开来，并阐明集成方法能够和不能表示什么（几何、归因子空间和秩限制）。

我们认为，GP/CGP 为线性高斯推断提供了一种以表示为先的框架，其中卡尔曼滤波器及其集成变体是高斯条件化的计算实现。Ens-CGP 是通过将经验集成矩视为一个（通常为低秩的）高斯先验，并在 $(\boldsymbol{H}, \boldsymbol{R})$ 下进行条件化而得到的、由集合实例化的条件高斯分布。这阐明了表示的内容（先验分布及其条件分布）与算法层面的内容（集合卡尔曼滤波变体、局地化、膨胀和迭代反演方案）之间的区别，并明确指出反复使用固定数据何时会偏离贝叶斯条件化。该框架与再生核希尔伯特空间及几何平方根表示之间的紧密联系，使其能够应用于逆问题、数据同化和基于核的回归。

致谢

本研究隶属于MIT气候大挑战、贾米尔天文台 CREWSNet以及天气与气候极端项目。Schmidt Sciences, LLC与曼谷银行慷慨支持了本项目。作者使用OpenAI ChatGPT作为写作助手，以润色语言并就表述提出修改建议。作者亦使用Grammarly进行编辑。所有技术内容、论断与错误仍由作者负责。

参考文献

[1] Robert J. Adler. 3. 样本函数正则性. 载于随机场几何，应用数学经典，第39–65页。工业与应用数学学会，2010年1月。[2] Robert J. Adler与Jonathan E. Taylor. 高斯场. 载于随机场与几何，第7–48页。Springer New York，纽约，NY，2007年。[3]Lubna AlBarghouty、Sai Ravela、William L. Rodi与Delaine T. Reiter. 面向多事件定位与震相关联的Transformer加速信息变分推断. 载于*AGU*秋季会议摘要，*AGU*秋季会议摘要2024卷，第S11A–04页，2024年12月。[4] Brian D.O. Anderson与John B. Moore. 离散时间卡尔曼滤波器. 载于最优滤波。Dover Publications，1979年。[5] Jeffrey L. Anderson. 一种用于数据同化的集合调整卡尔曼滤波器. 《月刊天气评论》，129(12):2884–2903，2001年。[6] Nathan Aronszajn. 再生核理论. 《美国数学学会汇刊》，68:337–404，1950年。[7] Alain Berlinet与Christine Thomas-Agnan. 再生核希尔伯特空间与随机过程. 载于概率与统计中的再生核希尔伯特空间，第55–108页。Springer US，马萨诸塞州波士顿，2004年。[8] Patrick Billingsley. 随机过程. 载于概率与测度，*Wiley*概率与统计丛书第939卷。John Wiley & Sons, Inc.，美国纽瓦克，第4版，2012年。[9] Christopher M. Bishop. 分类线性模型. 载于模式识别与机器学习，信息科学与统计学。Springer，纽约，2006年。[10] Jon Cockayne、Chris J. Oates、T. J. Sullivan与Mark Girolami. 贝叶斯概率数值方法. 《*SIAM*评论》，61(4):756–789，2019年1月。[11] Noel Cressie. 空间预测与克里金法. 载于空间数据统计学，Wiley概率与统计丛书，第105–209页。John Wiley & Sons, Inc.，1993年9月。[12] Geir Evensen. 集合卡尔曼滤波器：理论表述与实际实现. 《海洋动力学》，53(4):343–367，2003年11月。[13] Alfredo Garbuno-Inigo、Franca Hoffmann、Wuchen Li与Andrew M. Stuart. 相互作用朗之万扩散：梯度结构与集合卡尔曼采样器. 《*SIAM*应用动力系统杂志》，19(1):412–441，2020年1月。

[14] Kenneth Gee 和 Sai Ravela。利用再分析数据和观测数据学习热带气旋增强的随机微分方程模型，2026。[15] Roger G. Ghanem 和 Pol D. Spanos。随机过程的表示。载于 Roger G. Ghanem 和 Pol D. Spanos 编，《style id='3'>随机有限元：谱方法，第 15–65 页。Springer New York，纽约州纽约市，1991。[16] Thomas M. Hamill、Jeffrey S. Whitaker 和 Chris Snyder。集合卡尔曼滤波器中背景误差协方差估计的距离依赖滤波。《月刊天气评论》，129(11):2776–2790，2001。[17] Thomas Hofmann、Bernhard Sch¨olkopf 和 Alexander J. Smola。机器学习中的核方法。《统计年刊》，36(3):1171–1220，2008。[18] Marco A I glesias、Kody J H Law 和 Andrew M Stuart。用于逆问题的集合卡尔曼方法。《逆问题》，29(4):045001，2013年3月。[19] Ian T. Jolliffe 和 Jorge Cadima。主成分分析：回顾与最新进展。《英国皇家学会哲学汇刊*A*辑：数学、物理与工程科学》，374(2065):20150202，2016年4月。[20] R. E. Kalman。线性滤波与预测问题的新方法。《基础工程学报》，82(1):35–45，1960年3月。[21] Jae Deok Kim、Sai Ravela 和 Rob L. Evans。将神经物理反演器应用于现场数据：用于隐伏地热系统地下电阻率估算的物理耦合机器学习框架。载于*AGU*秋季会议摘要，第 2025 卷，第 NG23A–02 页，2025 年 12 月。[22] George S. Kimeldorf 和 Grace Wahba。随机过程的贝叶斯估计与样条平滑之间的对应关系。《数学统计年刊》，41(2):495–502，1970。[23] David M. Livings、Sarah L. Dance 和 Nancy K. Nichols。无偏集合平方根滤波器。《物理学*D*辑：非线性现象》，237(8):1021–1028，2008年6月。[24] David J. C. MacKay。用于主动数据选择的基于信息的目标函数。《神经计算》，4(4):590–604，1992年7月。[25] Georges Matheron。地质统计学原理。《经济地质学》，58(8):1246–1266，1963年12月。[26] Marzieh Alireza Mirhoseini。用少量数据学习降雨驱动的淹没替代模型，2024。

[27] A. O’Hagan。用于预测的曲线拟合与最优设计。皇家统计学会期刊。*B*辑（方法学），40(1):1–42，1978。[28] Carl Edward Rasmussen 和 Christopher K. I. Williams。协方差函数。载于机器学习中的高斯过程，第 0 页。麻省理工学院出版社，2005年11月。[29] Carl Edward Rasmussen 和 Christopher K. I. Williams。回归。载于机器学习中的高斯过程，第 0 页。麻省理工学院出版社，2005年11月。[30] Michael Reed 和 Barry Simon。II——希尔伯特空间。载于现代数理物理方法，第 36–66 页。学术出版社，1972年1月。[31] Anamitra Saha 和 Sai Ravela。利用数据与模型对极端降雨进行统计物理对抗学习和降尺度。地球系统建模进展期刊，16(6):e2023MS003860，2024年6月。[32] Simo S¨arkk¨a。贝叶斯推断。载于贝叶斯滤波与平滑，《数理统计学会教材》，第 17–26 页。剑桥大学出版社，英国剑桥，2013。[33] Claudia Schillings 和 Andrew M. Stuart。用于逆问题的集合卡尔曼滤波器分析。*SIAM*数值分析期刊，55(3):1264–1290，2017年1月。[34] Bernhard Sch¨olkopf 和 Alexander J. Smola。核。载于核方法学习：支持向量机、正则化、优化及其他，第 25–60 页。麻省理工学院出版社，2001。[35] Bernhard Sch¨olkopf 和 Alexander J. Smola。正则化。载于核方法学习：支-持向量机、正则化、优化及其他，第 87–124 页。麻省理工学院出版社，2001。[36] Ingo Steinwart 和 Andreas Christmann。核与再生核希尔伯特空间。载于支持向量机，第 110–163 页。Springer New York，纽约州纽约市，2008。[37] Albert Tarantola。一般离散反问题。载于反问题理论与模型参数估计方法，《应用数学其他著作》，第 1–40 页。工业与应用数学学会，2005。[38] Albert Tarantola。最小二乘准则。载于反问题理论与模型参数估计方法，《应用数学其他著作》，第 57–80 页。工业与应用数学学会，2005。[39] Michael K. Tippett、Jeffrey L. Anderson、Craig H. Bishop、Thomas M. Hamill 和 Jeffrey S. Whitaker。集合平方根滤波器。《月刊天气评论》，131(7):1485–1490，2003。[40] Grace Wahba。更多样条。载于观测数据样条模型，《CBMS-NSF应用数学区域会议系列》，第 21–39 页。工业与应用数学学会，1990年1月。