同步函数分位数回归的局部稀疏估计

Boyi Hu，西蒙菲莎大学统计与精算科学系；Jiguo Cao，西蒙菲莎大学统计与精算科学系

arXiv:2602.01691v1 [stat.ME] 2 Feb 2026

摘要

受探究日气温如何影响大豆产量这一问题的启发，本文提出一种同步函数型分位数回归（FQR）模型，其独特之处在于引入了一个同时由分位数和时间索引、并与函数型预测变量相关联的局部稀疏二元斜率函数。斜率函数的局部稀疏性意味着它仅在定义域的某些区段取非零值，在其他区段则为零。这些随分位数变化的零斜率区域表明，在相应时段，函数型预测变量对响应变量没有明显影响。这一特性增强了模型的可解释性。与一次拟合一个分位数且存在若干局限的传统FQR模型不同，我们提出的方法能够同时处理一系列分位数。我们通过模拟研究检验了这一新方法，结果显示其明显优于标准方法。为验证其实际应用价值，我们将该方法用于大豆产量数据，确定了日气温不会影响产量的时段。这一发现对农业规划和作物管理可能至关重要。

关键词：函数型数据分析；函数型分位数回归；二元样条；稀疏估计。

1 引言

作物产量在农业中发挥着关键作用，对满足全球对食品、饲料和燃料的需求具有深远影响。在全球主要作物中，大豆尤为突出，尤其是因为其产量的大部分——超过四分之三——被用作牲畜饲料。比例小得多的部分，约7%，则用于豆腐和豆浆等传统大豆制品。鉴于这一背景，了解环境因素如何影响大豆产量，对于确保稳定且可持续的农业实践至关重要。

大豆是一种对环境条件要求较高的作物，尤其是对温度和水资源可用性（Schlenker 和 Roberts，2009年；Vogel 等人，2019年）。大豆萌发和生长存在一个最佳温度范围，在此范围内，小幅波动不会对植物发育或最终产量造成不利影响。本项目旨在探究温度和水资源可用性如何影响生长季末的大豆产量。此外，我们还希望确定温度适宜大豆生长的具体时期，以及温度过低或过高、可能阻碍植物发育的时期。

为解决这些问题，我们提出一种新的函数型分位数回归模型，以研究大豆产量条件分位数与所选环境因素之间的关系，

$$Q_Y(u|\boldsymbol{Z}, X) = \boldsymbol{Z}^\tau \boldsymbol{\alpha}(u) + \int_0^T \beta(t,u)X(t)dt, \tag{1}$$

其中，Y是标量响应，$\boldsymbol{Z} = (1, Z_1, \ldots, Z_p)^\tau$是由截距和标量预测变量构成的向量，$X(t)$是在 $[0,T]$上定义的函数型预测变量，$Q_Y(u|\boldsymbol{Z}, X)$是 Y的条件分位数函数，u 是分位数，而 $\boldsymbol{\alpha}(u) = (\alpha_0(u) \ldots, \alpha_p(u))^\tau$和 $\beta(t,u)$是未知的斜率函数。通过关注分位数特异性效应，我们可以确定温度对产量结果产生显著影响的时间段，获得传统基于均值的分析无法揭示的信息。例如，在我们的大豆产量研究中，Y 表示年度大豆产量，这是我们感兴趣并希望进行建模和预测的响应变量，

而 $X(t)$表示感兴趣时间区间内的日平均气温，它是一个随时间变化的函数型预测变量 t。我们使用标量预测变量来刻画各县的水资源可用性。具体而言， Z_1 是各县的年降水量， Z_2 是县级用于大豆种植的灌溉土地比例。对于参数，向量 $\boldsymbol{\alpha}(u)$的每个分量对应一个变系数，用于刻画截距或标量预测变量对响应变量 Y的 u分位数的影响。变系数 $\boldsymbol{\alpha}(u)$使我们能够捕捉标量预测变量的影响如何随大豆产量分布的不同分位数而变化。函数 $\beta(t,u)$是一个以时间 t 和分位数 u为索引的二元斜率函数。它描述函数型预测变量 $X(t)$对大豆产量 Y各分位数的动态影响。

由于大豆生长存在一个适宜的温度范围，在此范围内，温度的小幅波动对大豆产量影响甚微，因此在模型 (1) 下，当时间 t 的温度处于这一适宜范围内时， $\beta(t,\ u)$ 应等于 0。换言之，对于固定的分位数 u，存在一些时间区域 $\mathcal{M} \subset [0,T]$，使得 $\beta(t,\ u) = 0$ （ $t \in \mathcal{M}$），其中 $\mathcal{M}$ 可能随分位数 u 而变化。这一观察强有力地促使我们在函数型分位数回归模型 (1) 中，为函数系数 $\beta(t,\ u)$ 引入局部稀疏性的概念。

局部稀疏性指的是在 $\beta(t,\ u)$ 的定义域内存在一些子区域，在这些区域中 $\beta(t,\ u) = 0$。换言之，在生长季的某些特定时间区间（子区域）内，日平均气温不会影响大豆产量的某些分位数。这些子区域对应于大豆生长的适宜温度范围，在该范围内，温度波动不会显著影响作物在某些特定分位数下的表现。更形式化地，我们假设存在一组子区域 $\mathcal{N} \subset [0,T] \times [0,1]$，使得对于所有 $(t,\ u) \in \mathcal{N}$， $\beta(t,\ u) = 0$。

通过纳入　　将 $\beta(t,\ u)$ 的局部稀疏性纳入模型后，我们便能更准确地

估计大豆产量分位数与日平均气温之间的动态依赖关系。具体而言，在模型（1）中识别 $\beta(t,\ u)$ 的局部稀疏区域，可以使模型聚焦于温度对大豆产量有显著影响的时段，同时忽略温度波动没有影响的子区域。此外，这也能更准确、更有意义地揭示温度如何影响不同分位数下的大豆产量。

函数型分位数回归（FQR）是数据分析中的重要工具，尤其适用于环境科学相关应用。它扩展了经典分位数回归（Koenker 和 Bassett Jr，1978年），能够对涉及函数型预测变量的数据进行更加灵活和全面的建模。研究人员针对不同情形并适应多样化的数据特征，发展了函数型分位数回归的各种扩展。例如，Cardot 等人（2005年）、Chen 和 M¨uller（2012年）以及 Kato（2012年）研究了仅包含单个函数型协变量且不含非函数型协变量的函数型分位数回归模型。Lu 等人（2014年）提出了函数型部分线性分位数回归模型，该模型纳入了有限维标量协变量。Yu 等人（2016年）考虑了包含多个函数型协变量和有限个标量协变量的模型。Yao 等人（2017年）提出了一个部分函数型分位数回归模型，其中包含一个函数型协变量和高维标量协变量。Qingguo 和 Kong（2017年）研究了半参数函数型分位数回归模型，其中标量协变量同时包含参数部分和半参数部分。Ma 等人（2019年）提出了一种函数型部分线性模型，可处理多个函数型协变量和超高维标量协变量。最近，Beyaztas 等人（2024年）提出了一种针对标量对函数分位数回归的稳健估计程序，从而能够在存在离群值和杠杆点的情况下实现可靠预测。值得注意的是，所有这些方法都采用传统的估计策略，即针对不同的分位数水平分别拟合模型。

具体而言，t 拟合函数型分位数回-的传统估计策略

归模型（1）的做法是，先固定分位数 u ，然后将 $\boldsymbol{\alpha}(u)$的每一项估计为标量参数，并将 $\beta(t,u)$估计为仅关于 t 的一元函数。虽然这种估计方法具有降低参数空间维数的优点，但也存在一个显著缺点：在估计二元函数 $\beta(t,u)$时缺乏光滑性。当针对不同分位数分别拟合模型时，所得 $\beta(t,u)$的估计量在时间域上关于t可能呈现光滑性，但无法保证在涉及分位数 u的方向上具有光滑性。这种光滑性的缺失会导致所估计的 $\beta(t,u)$在各分位数之间表现不稳定，从而可能无法准确刻画函数型预测变量与响应变量之间真实的潜在关系。

在最近的一项研究中，Hu等人（2024）针对仅包含单个函数型协变量且不含非函数型协变量的函数型分位数回归（FQR），提出了一种同时估计策略。他们将对观测到的函数型协变量应用函数主成分分析（FPCA）得到的函数主成分作为基函数，用于逼近斜率函数。随后，该模型在多个分位数水平上进行联合拟合，而不是分别拟合。尽管Hu等人（2024）为FQR引入了一种新颖的估计框架，但该模型仍相对简单，其可解释性也受到限制。在实践中，结果变量可能并非在整个时间域或所有分位数水平上都与函数型协变量相关。因此，识别存在显著关联的具体时间区域和分位数水平至关重要。然而，Hu等人（2024）并未提供用于检测此类区域的推断方法。对于未知二元函数进行推断可能颇具挑战性，例如同时依赖于时间 t 和分位数指标 u的斜率函数 $\beta(t,\ u)$。另一种策略是着重识别不存在关联的区域，从而缩小潜在显著区域的定义域。这推动了针对函数回归模型中未知斜率函数的局部稀疏估计研究。局部稀疏估计是指

在某些子区域中以稀疏性估计斜率函数的策略——这些子区域对应于函数型协变量对结果变量没有影响的时间—分位数域。

目前，函数型数据分析领域中探索局部稀疏性的研究大多聚焦于函数型线性模型（Reiss et al., 2017），这类模型已在多种应用中展现出实用价值（James et al., 2009; Zhou et al., 2013）。将局部稀疏估计融入这些模型，可以进一步提升其可解释性（Lin et al., 2017; Fang et al., 2020; Li et al., 2022; Gurer et al., 2024）。具体而言，Lin et al. (2017) 将 Fan and Li (2001) 提出的平滑截断绝对偏差（SCAD）扩展到函数型情形，称为 fSCAD。该扩展旨在为单变量标量对函数回归构造局部稀疏估计量。在此基础上，Fang et al. (2020) 和 Li et al. (2022) 随后将 fSCAD 用于处理函数型线性回归中的多个输出。在函数型分位数回归（FQR）领域，Yu 等人 (2019) 和 Liang et al. (2023) 探讨了使用不同正则化工具诱导斜率函数稀疏性的局部稀疏估计。然而，这些现有方法沿用传统的估计策略：固定分位数水平，并将斜率函数估计为时间的一元函数。这严重限制了恢复结果的条件分布与函数型协变量之间平滑关系的能力。

在我们提出的方法中，我们使用基于三角剖分的二元样条来近似感兴趣的非参数成分 $\beta(t,\ u)$。在此基础上，我们开发了一种新方法来估计 $\beta(t,\ u)$，该方法能够控制光滑性，并保证 $\beta(t,\ u)$ 的局部稀疏性。据我们所知，本文首次尝试为函数型分位数回归开发一种兼具局部稀疏性与光滑性的估计方法。我们的方法革新了传统的估计策略，在模型拟合过程中同时纳入多个分位数，克服了只能固定单个分位数的局限。

为实现这些创新，我们的方法采用粗糙度正则化，控制将 $\beta(t,\ u)$ 视为二元函数时其估计的光滑性。此外，基于三角剖分上定义的分片二元样条的性质，我们能够在三角形层级对 $\beta(t,\ u)$ 的近似施加惩罚项，从而实现局部稀疏估计。将光滑性控制与局部稀疏正则化结合使用，不仅增强了模型的可解释性，也提高了 $\beta(t,\ u)$ 的估计精度。本文接下来的数值研究将展示这些改进。

本文其余部分安排如下。第 2 节介绍模型及其对应的 $\beta(t,u)$ 估计量。第 3 节给出主要理论结果。第 4 节首先介绍一个大豆产量数据集，然后将所提方法应用于该数据集进行分析，以作说明。第 5 节总结全文。用于复现文中模拟研究和应用的计算代码可在 `https://github.com/caojiguo/LocalFQR/`获取。

2 同步函数分位数回归

对于所提出的具有局部稀疏斜率函数 $\beta(t,u)$ 的函数型分位数回归模型 (1)，我们用 $\mathscr{U} \subset (0,1)$ 表示包含关注分位数的区间，并用 $\Omega = [0,T]\times\mathscr{U}$ 表示所关注的 $\beta(t,u)$ 的定义域。令 $S_0 = \{(t,u) \in \Omega : \beta(t,u) = 0\}$ˆ 表示 $\beta(t,u)$ 的零区域。令 S_0 表示利用数据识别出的 S_0。下一步是估计补集 $\Omega\backslash S_0$ 上的函数 $\beta(t,u)$。该补集由函数 $\beta(t,u)$ 非零的区域构成，这些区域正是我们关注的部分。

直接估计无限维函数 $\boldsymbol{\alpha}(u)$ 和 $\beta(t,u)$ 在实践中可能颇具挑战。为解决这一问题，一种常见策略是使用基函数的线性组合来近似这些未知函数，然后根据数据估计这些基函数的系数。

假设 $\beta(t,u)$ 足够平滑，我们提出使用双变量样条基函数近似 $\beta(t,u)$。关于二元样条基，有多种可用于近似的方法。两种常用方法是 B样条张量积（Stone 等，1997）和定义在三角剖分上的双变量伯恩斯坦多项式（Lai 和 Schumaker，2007年；Lai 和 Wang，2013；Li 等，2021年）。本文选择在三角剖分上定义的双变量伯恩斯坦多项式，来近似式（1）中的未知斜率函数 $\beta(t,u)$。每个双变量伯恩斯坦多项式的支撑位于三角剖分中的单个三角形上，并且每个三角形上都定义了多个伯恩斯坦多项式。这种局部支撑性质确保我们可以利用这些三角形来估计子集 S_0。

令 $S_d^r(\Delta)$ 表示定义在三角剖分 Δ 上的二元样条所张成的线性空间，其中 r 是该空间的光滑性条件， d 是样条的次数。我们用 $|\Delta|$ 表示三角剖分 Δ 中三角形的最长边。我们的目标是找到一个函数 $s(t,u) \in S_d^r(\Delta)$，使其能够有效地逼近定义域 Ω 上的斜率函数 $\beta(t,u)$。为了使后续章节的叙述和证明更加清晰，我们用 $\{B_j(t,u)\}_{j=1}^{n_B}$ 表示定义在三角剖分$\Delta = \{\Lambda_1, \ldots, \Lambda_M\}$上的伯恩斯坦多项式，其中 $j = 1, \ldots, n_B$ 是多项式的索引。 n_B 与 M 之间的关系为 $n_B = (d+2)(d+1)M/2$ ，因为 $(d+2)(d+1)/2$ 伯恩斯坦多项式与 Δ 的每个三角形相关联。此外，对于每个二元基函数 $B_j(t,u)$，我们用 Δ_j 表示其支撑。换言之，当 $(t,u) \in \Delta_j$ 时， $B_j(t,u) \neq 0$；当 $(t,u) \notin \Delta_j$ 时， $B_j(t,u) = 0$。如果两个伯恩斯坦多项式 $B_j(t,u)$ 和 $B_k(t,u)$ 与同一个三角形相关联，那么 Δ_j 和 Δ_k 相同。补充文档中的图 S1 展示了 $d = 2$ 时定义在同一个三角形上的六个伯恩斯坦多项式。

近似函数 $s(t,u) \in S_d^r(\Delta)$ 可表示为伯恩斯坦多项式 $\{B_j(t,u)\}_{j=1}^{n_B}$ 的线性组合，它用于近似 $\beta(t,u)$。于是，在定义域 Ω 上，我们有

近似

$$\beta(t,u) \approx s(t,u) = \sum_{j=1}^{n_B} \gamma_j B_j(t,u) := \boldsymbol{B}(t,u)^\tau \boldsymbol{\gamma} \in S_d^r(\Delta), \tag{2}$$

其中，$\boldsymbol{B}(t,u) = (B_1(t,u)\ldots, B_{n_B}(t,u))^\tau$是定义在 Ω 上的双变量样条基函数向量，而 $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{n_B})^\tau$是相应的基系数向量。

为了确保 $\beta(t,u)$ 的近似具有所需的光滑性，例如函数连续且导数连续，需要对基系数 $\boldsymbol{\gamma}$施加线性约束。这些约束可以表示为 $\boldsymbol{H}\boldsymbol{\gamma} = 0$，其中 $\boldsymbol{H}$ 是线性约束矩阵。通过纳入这些约束，我们可以在估计中保证所需的光滑性。然而，对系数施加线性约束可能会使优化过程变得复杂。为解决这一问题，一种常见做法是使用 QR分解消除线性约束；在估计广义加性模型时，这也是处理可识别性约束的常用技术（Wood，2017年）。对于给定的 $\boldsymbol{H}$，通过 QR分解，我们有

$$\boldsymbol{H}^{\mathrm{T}} = (\boldsymbol{Q}^*, \boldsymbol{Q}) \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix}, \tag{3}$$

其中（$\boldsymbol{Q}^*, \boldsymbol{Q}$）是一个具有正交列的矩阵，且 $\boldsymbol{R}$ 是一个对角元素非零的上三角矩阵。基于分解（3），可通过将 $\boldsymbol{\gamma}$ 改写为 $\boldsymbol{\gamma} = \boldsymbol{Q}\boldsymbol{\theta}$来去除约束 $\boldsymbol{H}\boldsymbol{\gamma} = \boldsymbol{0}$，其中 $\boldsymbol{\theta}$ 是一个维度低于原始参数向量 $\boldsymbol{\gamma}$的向量。

当假设 $\boldsymbol{\alpha}(u)$ 中的一元函数足够平滑时，$\boldsymbol{\alpha}(u)$ 的各项可表示为具有未知系数的 B样条基函数的线性组合：$\alpha_k(u) = \sum_{j=1}^{n_b} \eta_{k,j} b_j(u) := \boldsymbol{b}(u)^\tau \boldsymbol{\eta}_k$，其中 $\boldsymbol{b}(u) = (b_1(u)\ldots, b_{n_b}(u))^\tau$是 B样条基函数向量（我们用 n_b 表示 B样条基函数的数量），定义在 $\mathscr{U}$上，而 $\boldsymbol{\eta}_k = (\eta_{k,1}, \ldots, \eta_{k,n_b})^\tau$是相应的基系数向量。于是，函数型分位数回归模型 (1) 可近似为 $Q_Y(u|X, \boldsymbol{Z}) \approx \boldsymbol{Z}^\tau \boldsymbol{V}(u)\boldsymbol{\eta} + \int \boldsymbol{B}(t,u)^\tau \boldsymbol{Q}\boldsymbol{\theta} X(t)dt \approx \boldsymbol{Z}^\tau \boldsymbol{V}(u)\boldsymbol{\eta} + \boldsymbol{A}(u)^\tau \boldsymbol{Q}\boldsymbol{\theta}$,

其中 $\boldsymbol{V}(u)$ 是由 B样条基函数构成的分块对角矩阵， $\boldsymbol{\eta}=(\boldsymbol{\eta}_0^\tau$，……， $\boldsymbol{\eta}_p^\tau)^\tau$和$\boldsymbol{A}(u)^\tau=\int \boldsymbol{B}(t,\ u)^\tau X(t)dt$。

在所提出的估计过程中，目标是同时考虑所有感兴趣的分位数，以估计单变量函数$\boldsymbol{\alpha}(u)$和二元函数 $\beta(t,\ u)$。在估计过程中结合多个分位数回归模型的优势已得到广泛认可，例如 Zou 和 Yuan（2008）以及 He 等（2016）的研究。

对于实值随机变量 Y， $E\{\rho_u(Y-u)\}$ 的最小化解是 Y的 u分位数，其中 $\rho_u(x)=x(u-\mathbb{1}\{x<0\})$称为分位数损失函数或检验函数（Koenker 和 Bassett Jr，1978年）。假设我们观测到独立同分布数据对 $\{y_i,\boldsymbol{z}_i,x_i(t)\}_{i=1}^n$，它们是 $\{Y,\boldsymbol{Z},X(t)\}$的实现。为了进行分位数回归，我们考虑一组感兴趣的分位数，记为 $U\in\mathscr{U}$。假设这些分位数在区间 $\mathscr{U}$ 内均匀分布。 U的基数记为 n_u，表示感兴趣的分位数个数。我们拟通过最小化下列损失函数来估计基系数，其中优化变量为 $\boldsymbol{\eta}$ 和 $\boldsymbol{\theta}$:

$$\frac{1}{nn_u}\sum_{r=1}^{n_u}\sum_{i=1}^{n}\rho_{u_r}\left\{y_i-\boldsymbol{z}_i^\tau\boldsymbol{V}(u_r)\boldsymbol{\eta}-\boldsymbol{A}_i(u_r)^\tau\boldsymbol{Q}\boldsymbol{\theta}\right\} \tag{4}$$

其中， $\boldsymbol{A}_i(u_r)=\int \boldsymbol{B}(t,\ u_r)^\tau x_i(t)dt$。

然而，式（4）中涉及 $\boldsymbol{z}_i$、 $\boldsymbol{b}$（u_r）和 A_i（u_r）的设计矩阵可能存在病态，直接最小化式（4）会导致 β（$t,\ u$）的估计量高度波动。因此，对 β（$t,\ u$）的估计可能不如模型假设的那样光滑。为了解决这一问题，一种常见的方法是在最小化过程中向式（4）添加惩罚项。该惩罚项可以提高数值稳定性，并控制 β（$t,\ u$）估计的光滑性。粗糙度惩罚是实现这一目的的一种特别有效的工具。关于函数型数据分析中粗糙度惩罚的详细讨论，可参见 Cardot 等（2003）、Ramsay 和 Silverman（2005）、Sangalli 等（2013）以及 Azzimonti 等（2015）。对于光滑的二元函数 s（$t,\ u$），粗糙度惩罚

R（s）定义为

$$R(s) = \sum_{\Lambda \in \Delta} \int_{\Lambda} \sum_{d_1+d_2=2} \binom{2}{d_1} \left[\nabla_t^{d_1} \nabla_u^{d_2} s(t,u)\right]^2 dtdu. \tag{5}$$

对于任意二元样条逼近 $s(t,u) = \boldsymbol{B}(t,u)^\tau \boldsymbol{Q}\boldsymbol{\theta}$，$s(t,u)$ 的粗糙度惩罚 (5) 可进一步写成带有半正定矩阵 $\boldsymbol{G}$ 的二次型 $R(s) = \boldsymbol{\theta}^\tau \boldsymbol{G}\boldsymbol{\theta}$。借助粗糙度惩罚 $R(\cdot)$，我们提出通过最小化以下目标函数来估计 (1) 中未知的系数 $\boldsymbol{\eta}$ 和 $\boldsymbol{\theta}$：

$$\frac{1}{nn_u} \sum_{r=1}^{n_u} \sum_{i=1}^{n} \rho_{u_r} \left(y_i - \boldsymbol{z}_i^\tau \boldsymbol{V}(u_r)\boldsymbol{\eta} - \boldsymbol{A}_i(u_r)^\tau \boldsymbol{Q}\boldsymbol{\theta}\right) + \lambda \boldsymbol{\theta}^\tau \boldsymbol{G}\boldsymbol{\theta}. \tag{6}$$

2.1 $\beta(t,u)$ 的局部稀疏估计

我们的主要目标是获得 $\beta(t,u)$ 的局部稀疏估计量，以帮助识别函数型预测变量 $X(t)$ 在标量响应 Y 的不同分位数下的非活动区域。为实现这一目标，我们采用组LASSO方法。

利用基于三角剖分的样条逼近方法，我们可以在三角形层级对逼近 (2) 的 L_2-范数施加惩罚，从而促进目标函数 $\beta(t,u)$ 在三角形层级上的局部稀疏性。为此，我们提出在估计过程中加入组 LASSO 惩罚项。实质上，组 LASSO 惩罚项会促使某些三角形上的二元样条系数为零，从而有助于识别函数型预测变量 $X(t)$ 在 Y 特定分位数下的非活动区域。具体而言，我们提出以下双重惩罚目标函数，以获得 $\beta(t,u)$ 的平滑且局部稀疏估计量：

$$\frac{1}{nn_u} \sum_{r=1}^{n_u} \sum_{i=1}^{n} \rho_{u_r} (y_i - \boldsymbol{z}_i^\tau \boldsymbol{V}(u_r)\boldsymbol{\eta} - \boldsymbol{A}_i(u_r)^\tau \boldsymbol{Q}\boldsymbol{\theta}) + \lambda_1 \boldsymbol{\theta}^\tau \boldsymbol{G}\boldsymbol{\theta} + \lambda_2 \sum_{j=1}^{J} \|\boldsymbol{\gamma}_{[j]}\|_2, \tag{7}$$

它结合了粗糙度惩罚和组 LASSO 惩罚。由此得到的 $\beta(t,\ u)$ 基于双变量样条的估计量具有平滑性，同时在三角形（组）层面呈现稀疏性。与标准套索回归一样，组 LASSO 惩罚可能会引入收缩偏差

在参数估计中。因此在实际应用中，我们考虑一种自适应版本，根据初始估计量赋予各组特定的权重，例如式 (6) 的极小化解。自适应组 LASSO 权重的选择并非唯一，可以采用几种合理的加权方案。本文考虑两种不同类型的组 LASSO˜惩罚。下文将式 (6) 的极小化解记为 $(\tilde{\boldsymbol{\eta}},\ \boldsymbol{\theta})$。

选项 1. 定义 $w_j = \left\|\tilde{\boldsymbol{\gamma}}_{[j]}\right\|_2^{-\eta}$，其中 $\tilde{\boldsymbol{\gamma}} = \boldsymbol{Q}\tilde{\boldsymbol{\theta}}$ 和 $\tilde{\boldsymbol{\gamma}}_{[j]}$ 表示与三角形 Δ_j对应的 $\tilde{\boldsymbol{\gamma}}$ 的各项。我们可以最小化以下目标函数，以获得 $\boldsymbol{\eta}$ 和 $\boldsymbol{\theta}$的估计值，

$$\frac{1}{nn_u}\sum_{r=1}^{n_u}\sum_{i=1}^{n}\rho_{u_r}(y_i - \boldsymbol{z}_i^\tau \boldsymbol{V}(u_r)\boldsymbol{\eta} - \boldsymbol{A}_i(u_r)^\tau \boldsymbol{Q}\boldsymbol{\theta}) + \lambda_1\boldsymbol{\theta}^\tau\boldsymbol{G}\boldsymbol{\theta} + \lambda_2\sum_{j=1}^{J} w_j \left\|\boldsymbol{\gamma}_{[j]}\right\|_2, \tag{8}$$

其中$\boldsymbol{\gamma}_{[j]}$ 表示与三角形 Δ_j对应的 $\boldsymbol{\gamma}$ 的各项。

˜**选项 2.** 将 $\beta(t,u)$ 的初始估计值记为 $\beta(t,u) = \boldsymbol{B}(t,u)^\tau\tilde{\boldsymbol{\gamma}}$。定义 $w_j = \left\|\tilde{\beta}|_{\Delta_j}\right\|_2^{-\eta}$，其中 $\tilde{\beta}|_{\Delta_j}$是 $\tilde{\beta}(t,u)$ 在三角形 Δ_j上的限制。另一种获得 $(\boldsymbol{\eta},\boldsymbol{\theta})$ 合理估计值的方法是最小化以下目标函数，

$$\frac{1}{nn_u}\sum_{r=1}^{n_u}\sum_{i=1}^{n}\rho_{u_r}(y_i - \boldsymbol{z}_i^\tau \boldsymbol{V}(u_r)\boldsymbol{\eta} - \boldsymbol{A}_i(u_r)^\tau \boldsymbol{Q}\boldsymbol{\theta}) + \lambda_1\boldsymbol{\theta}^\tau\boldsymbol{G}\boldsymbol{\theta} + \lambda_2\sum_{j=1}^{J} w_j \left\|f_{\boldsymbol{\theta}}|_{\Delta_j}\right\|_2, \tag{9}$$

其中，$f_{\boldsymbol{\theta}}(t,\ u) = \boldsymbol{B}(t,\ u)^\tau\boldsymbol{Q}\boldsymbol{\theta}$，且 $f_{\boldsymbol{\theta}}|_{\Delta_j}$是 $f_{\boldsymbol{\theta}}$在 Δ_j上的限制。

对于组 LASSO 惩罚项的两种形式，要最小化的相应目标函数由多个分位数水平上的分位数损失之和以及两个正则化项组成。第一个正则化项是以二次型表示的粗糙度惩罚，用于控制估计的二元斜率函数的光滑性。第二个是组 LASSO 惩罚项，用于促进估计的斜率曲面的局部稀疏性。该最小化问题属于二阶锥规划（SOCP），可表述为符合规范凸规划（DCP）规则集的凸优化问题。因此，可使用 R 软件包 `CVXR` （Fu 等，2020）求解该问题，并可选择 `SCS` 或 `ECOS` 求解器。

分位数回归中的一个重要实际问题是避免条件分位数估计发生交叉，因为有效的条件分位数函数必须是非递减的

关于分位数指标 u （Hu et al., 2024；Neocleous and Portnoy，2008）。然而，当分别估计分位数或样本量有限时，通常无法保证估计的条件分位数函数 $Q_Y(u \mid X_i,\ \boldsymbol{Z}_i)$ 对所有个体 $i = 1, \ldots, n$都关于 u 非递减。文献中已提出多种方法来解决这一问题。Schnabel and Eilers（2013）考虑了这样一种情形：标量响应的条件分位数曲线由单个标量协变量索引。他们将分位数指标 u 视为额外的协变量，并提出将条件分位数函数估计为协变量和 u上的平滑曲面，同时施加非对称差分惩罚以促进分位数方向上的单调性。Takeuchi et al.（2006）提出同时估计多个分位数，并在观测数据点处施加显式的非交叉约束。Liu and Wu（2011）采用基于核的表示来表示分位数回归函数，并通过对核系数施加线性不等式约束来确保分位数不交叉。本文采用 Takeuchi et al.（2006）提出的一般思路，将单调性约束直接纳入式（8）和（9）中定义的优化程序。具体而言，我们可以先选择一组感兴趣的分位数，记为 S。然后施加不等式约束，以确保分位数之间的单调性：对于任意满足 $\tau_1,\ \tau_2 \in S$的 $\tau_1 > \tau_2$ ，每个个体 i的估计 τ_1分位数都不应小于估计的 τ_2分位数。这种方法有助于防止条件分位数估计发生交叉，从而保留真实分位数函数的单调性。然而，当样本量较大时，这种带约束的优化在计算上开销很大，因为约束数量会随着样本量和所考虑的分位数水平数量增加。此外，求解由此产生的问题需要在规范凸规划框架下使用商业求解器 `MOSEK` （ApS，2019）。开发一种更高效、易于使用的算法来处理此类约束，仍是一项尚未解决的挑战。基于这些原因，我们没有纳入非交叉分位数

将这一约束纳入当前框架，并将其留作未来研究的一个方向。

2.2 参数调优

在本小节中，我们介绍如何确定（8）或（9）中的调优参数 λ (6)、λ_1 和 λ_2 的取值。权重 w_j 对于组 LASSO 惩罚项至关重要，这些惩罚项用于在估计 $\beta(t,\ u)$ 时实现局部稀疏性，通常根据一些初始估计确定。在我们的框架中，我们建议采用与目标函数（6）相对应的估计方法，获得 $\boldsymbol{\theta}$的初始估计。该初始估计仅取决于调优参数 λ，且不具备任何局部稀疏特性，因为函数（6）中没有涉及稀疏正则化。在这一步中，我们选择使用 10 折交叉验证来确定 λ的最优取值，然后获得 $\boldsymbol{\theta}$的初始估计。基于 $\boldsymbol{\theta}$的初始估计，我们根据（8）或（9）计算权重 w_j 。

下一步是确定目标函数 (8) 或 (9) 中 λ_1 和 λ_2 的取值。为此，对于给定的一组调节参数，我们首先使用整个数据集，根据目标函数 (8) 或 (9) 拟合模型，从而从三角剖分中确定活动三角形集合。具体而言，我们根据通过最小化 (8) 或 (9) 得到的 $\hat{\boldsymbol{\gamma}}$，识别系数不为零的三角形，其中 $\hat{\boldsymbol{\gamma}} = \boldsymbol{Q\theta}$。随后，在每一轮交叉验证中，我们仅在活动三角形集合内重新拟合参数，并排除组 LASSO 惩罚项，以计算验证损失。这种方法有助于避免组 LASSO 惩罚在估计 $\boldsymbol{\gamma}$ 时引入的潜在偏差。

为确保在模型拟合中正确排除函数型预测变量的已识别非活动区域（系数为零的三角形），我们在估计过程中对 $\boldsymbol{\gamma}$ 施加额外的线性约束，以确保所选三角形的系数设为零。对 $\boldsymbol{\gamma}$ 施加的这些额外线性约束可表示为：$\boldsymbol{D\gamma} = \mathbf{0}$，其中矩阵 $\boldsymbol{D}$ 由 0 和 1 构成。这确保基

与这些系数为零的所选三角形相对应的基系数，在后续模型拟合步骤中仍保持为零。

在实践中，我们倾向于为粗糙度惩罚采用较小的候选值。尽管较小的值可能会增加斜率估计的变异性，并可能导致函数略显波动，但它们有助于减小估计偏差。因此，我们不建议使用非常大的值。当估计过程中同时纳入粗糙度惩罚和组 LASSO 惩罚时，这一点尤为重要。在这种情况下，较大的粗糙度惩罚可能导致估计结果过于稀疏，尤其是在信号区域与非信号区域的边界附近。由于这些边界处的曲率通常较大，较大的粗糙度惩罚可能会在这些区域造成过度惩罚，从而抑制真实信号并扭曲估计函数。

3 理论性质

定义

$$\Gamma_X f = \int_s \mathrm{cov}(s,t) f(s,u) ds, \quad \|f(t,u)\|_{\Gamma_X,2} = \left(\int_u \int_t \Gamma_X f \times f dt du \right)^{1/2}.$$

为推导通过最小化 (6) 得到的所提出估计量 $\beta(t,\ u)$ 的渐近性质，我们作出以下假设：

(A1) $\{Y_i, \boldsymbol{Z}_i, X_i(t)\}_{i=1}^n$ 是独立同分布的。

(A2) $\|X(t)\| \leq C_0 < \infty$ 几乎处处。

(A3) 假设 $\boldsymbol{\alpha}(u)$ 中的函数具有 p' 阶导数 $\boldsymbol{\alpha}^{(p')}(u)$，使得

$$\left| \alpha_k^{(p')}(t) - \alpha_k^{(p')}(s) \right| \leq C \left| t - s \right|^{\vartheta}, \quad s,t \in (0,1),$$

其中 $C_1 > 0$ 和 $\vartheta \in [0,1]$。接下来，我们令 $\nu = p' + \vartheta$。

(A4) Γ_X 的特征值均严格为正。

(A5) $\beta(t,\ u)\ \in W_q^{d+1}(\mathscr{T}\times\mathscr{U})$。

(A6) 由 $\epsilon_i = Y_i - \boldsymbol{Z}_i^{\intercal}\boldsymbol{\alpha}(u)\ -\int\beta(t,u)X_i(t)dt$ 定义的随机变量 ϵ_i 具有密度函数 f，该密度函数在0处连续，且下界为一个严格正的常数。

定理1 在假设*A1-A6*下，设 $n_b\ \sim\ |\Delta|^{-(d+1)/\nu}$，且 n_u 足够大，则

$$\|\hat{\beta}(t,u)-\beta(t,u)\|_{\Gamma_X,2} = O_p\left(|\Delta|^{d+1}+\frac{1}{n\lambda|\Delta|^2}+\lambda\right),$$

除了一个概率随 n 趋于无穷而趋于零的集合之外。

ˆ 接下来，我们为由(7)得到的 $\beta(t,u)$的 l_2 估计误差建立一个非渐近上界。不失一般性，我们假设(1)中的协变量分量 $\boldsymbol{Z}$仅由一个截距项组成。为记号简便，现在我们用 $\boldsymbol{\gamma}$ 同时表示与B样条和伯恩斯坦多项式相关联的参数向量。也就是说，$\boldsymbol{\gamma}$ 的前 n_b 个分量与B样条相关联，其余 n_B 个分量与三角剖分上的伯恩斯坦多项式相关联。采用这一记号，我们令 $\boldsymbol{\gamma}_0$ 表示前 n_b 个分量，$\boldsymbol{\gamma}_{-1}$ 表示其余 n_B 个分量。定义$\tilde{\boldsymbol{w}}_{i,u} = \left[\boldsymbol{z}_i^{\top}\boldsymbol{V}(u)\ \ \boldsymbol{A}_i(u)^{\top}\right]$，$\tilde{\boldsymbol{W}}_u = [\tilde{\boldsymbol{w}}_{i,u},\ldots,\tilde{\boldsymbol{w}}_{i,u}]^{\top}$，以及$\hat{\boldsymbol{\Sigma}} := n^{-1}\tilde{\boldsymbol{W}}_u^{\top}\tilde{\boldsymbol{W}}_u$。定义 $\kappa\in(0,1)$为

$$1-\kappa = \mathrm{P}\left(\|\hat{\boldsymbol{\Sigma}}_j - \boldsymbol{I}_{|G_1|}\|\| \le 0.5,\quad \forall 0\le j\le J,\quad \forall u\in U\right).$$

数值常数0.5并非必不可少，可以替换为任意固定的 $c\in(0,1)$。该条件定义了一个高概率事件，其补集的概率为 κ。当ˆκ 较小时，这意味着在 j 和 u上一致地，样本协方差矩阵 $\boldsymbol{\Sigma}_j$ 是良态的，并且保持接近其总体对应矩阵。我们进一步假设:

(A7)（受限特征值条件）。 $\phi_{\min} := \phi_{\min}(c_0,\bar{S}) := \inf_{\boldsymbol{a}\in\mathbb{C}\cap\mathbb{S}^{n_b+n_B-1}}\min_{u\in U}\left\|\boldsymbol{\Sigma}_u^{1/2}\boldsymbol{a}\right\| >0$，其中 $\boldsymbol{\Sigma}_u$ 是分位数 u 处进行样条逼近后设计矩阵的协方差矩阵，U 是选定的候选分位数集合。我们还定义 $\phi_{\max} :=\phi_{\max}(c_0,\bar{S}) := \sup_{\boldsymbol{a}\in\mathbb{C}\cap\mathbb{S}^{n_b+n_B-1}}\max_{u\in U}\left\|\boldsymbol{\Sigma}_u^{1/2}\boldsymbol{a}\right\|$。 $\phi_{\min}$ 和 $\phi_{\max}$ 均取决于三角剖分的构造。

取一个稀疏向量 $\bar{\boldsymbol{\gamma}}_{-1} \in \mathbb{R}^{n_B}$，使得 $\bar{\boldsymbol{\gamma}}_{-1}$ 满足线性约束 $\boldsymbol{H}\bar{\boldsymbol{\gamma}}_{-1} = \mathbf{0}$，且 $\beta(t,u)$ 的相应逼近误差被某个较小常数a_Δ所界定：

$$\sup_{(t,u)\in\Omega} |\bar{\beta}(t,u) - \beta(t,u)| \le a_\Delta.$$

此外，取另一向量 $\bar{\boldsymbol{\gamma}}_0$，使得截距函数 $\alpha(u)$ 的 B样条逼近误差也有界于 a_Δ。定义 $\bar{\boldsymbol{\gamma}}$ 为：$\bar{\boldsymbol{\gamma}} = (\bar{\boldsymbol{\gamma}}_0^\top, \bar{\boldsymbol{\gamma}}_{-1}^\top)^\top$。令 $C_{\bar{\beta}}$ 为粗糙度惩罚的值

$$\bar{\beta}(t,u) := \boldsymbol{B}(t,u)^\top \bar{\boldsymbol{\gamma}}_{-1}.$$

对于某个 $h > 0$，定义

$$r_* := 4C_f a_\Delta / c_f \phi_{\min} \vee h C_{\bar{\beta}},$$
$$\epsilon^*_{A_3} := \sqrt{3}\,(2\rho_{\max}(\boldsymbol{G})C_{\bar{\boldsymbol{\gamma}}}\lambda_1 + \lambda_2)\sqrt{J} + \frac{\sqrt{6}\,p_\Delta^{1/2} c_4^{1/2}\sqrt{J}}{\sqrt{n}} + \frac{\sqrt{6}\,p_\Delta^{1/2} c_4 A_3 \sqrt{J\log J}}{\sqrt{n}},$$

其中，$\rho_{\max}(\boldsymbol{G})$ 表示 $\boldsymbol{G}$ 的最大特征值。

定理2 假设条件 *(A1)-(A7)* 成立。令 A_1、A_2 和 A_3 为任意正数。我们选择调节参数 λ_1 和 λ_2，使得 $n_U\left((c_0-1)\lambda_2 - h^{-1}\lambda_1\right) > 2\sqrt{2n^{-1/2}} + 1.5C_f a_\Delta + A_1 n^{-1/2} + A_2\sqrt{\text{对数}}$ J/n。以至少 $1 - 2e^{-A_1^2} - 16Je^{-A_2^2/128} - 64J^{1-A_3^2} - \kappa$ 的概率，对于通过最小化 (7) 得到的 $\hat{\gamma}$，有

$$\|\hat{\boldsymbol{\gamma}} - \bar{\boldsymbol{\gamma}}\|_2 \le \frac{4}{c_f \phi_{\min}^2}\left[\epsilon^*_{A_3} \vee \left(\frac{24(c_0+1)\sqrt{2p_\Delta J}}{\sqrt{n}} \vee \frac{2\sqrt{2}\,\phi_{\max}}{\sqrt{n}}\right)\right] \vee r_*\,.$$

本节定理的详细证明见补充材料。

为了进一步理解误差界，我们考虑相应的渐近情形。在定理2中，假设调优参数 λ_1 和 λ_2 的选择满足 $2\rho_{\max}(\boldsymbol{G})C_{\bar{\boldsymbol{\gamma}}}\lambda_1 \lesssim n_u^{-1}n^{-1/2}\sqrt{\log}\ J$，且 $\lambda_2 \lesssim n_u^{-1}n^{-1/2}\sqrt{\log}\ J$。注意，$\rho_{\max}(\boldsymbol{G}) \asymp |\Delta|^{-2}$ 和 $C_{\bar{\boldsymbol{\gamma}}} \sim |\Delta|^{-1}$。在此选择下，$\epsilon^*_{A_3}$ 中的主导项为 $\sqrt{6}\,p_\Delta^{1/2} c_4 A_3 \sqrt{J\log}$ $J \asymp \sqrt{J\log}\ Jn^{-1/2}$。关于项 r_*，逼近误差的上界 a_Δ 的阶为

$O(|\Delta|^{d+1})$，三角形的数量为 $J \sim |\Delta|^{-2}$。若选择 $h \sim |\Delta|^{d+1}$，则$r_* \sim |\Delta|^{d+1}$。此外，我们作出以下假设：

(A8) $1 \lesssim \phi_{\min} \leq \phi_{\max} \lesssim 1$。

(A9) $\kappa \to \infty$ 为 $n \to \infty$。

由此可得以下推论。

推论1 假设条件 *(A1)-(A9)* 成立。取 $\lambda_1 \sim n_u^{-1} n^{-1/2}\sqrt{\text{对数}}\ J|\Delta|^3$，以及 $\lambda_2 \sim n_u^{-1} n^{-1/2}\sqrt{\text{对数}}\ J$，使得 $n_U((c_0-1)\lambda_2 - h^{-1}\lambda_1) > 2\sqrt{2}n^{-1/2} + 1.5C_f a_\Delta + A_1 n^{-1/2} + A_2\sqrt{\text{对数}\ J/n}$，其中 $h \sim |\Delta|^3$ 为某个常数。则当 $n \to \infty$ 时，我们有

$$\|\hat{\boldsymbol{\gamma}} - \bar{\boldsymbol{\gamma}}\|_2 = O_p\left(|\Delta|^{-1}(-2\log|\Delta|)^{-1/2} n^{-1/2} + |\Delta|^3\right).$$

根据 Lai和Schumaker（2007年），在推论1的条件下，对于某个常数 C，我们有

$$|\hat{\beta}(t,u) - \bar{\beta}(t,u)|_2 \leq C|\Delta|\|\hat{\boldsymbol{\gamma}} - \bar{\boldsymbol{\gamma}}\|_2 = O_p((-2\log|\Delta|)^{-1/2} n^{-1/2} + |\Delta|^4).$$

4 真实数据分析

温度在大豆各生长阶段（包括营养生长期和生殖生长期）都发挥着至关重要的作用。高温和低温都会对大豆萌发和生长产生不利影响。因此，了解不同分位数下日气温与大豆产量之间的关系，对于全面认识大豆生长动态至关重要。

本研究旨在控制其他环境因素的同时，考察日气温对不同分位数下大豆产量的影响。通过纳入年降水量和灌溉面积占比等因素，我们可以更好地了解温度对大豆产量的独立贡献。分位数回归使我们能够

分析分布中不同位置的变量关系，超越传统均值回归分析。通过考察多个分位数，我们可以确定温度在各种条件下（包括有利和不利情形）对大豆产量的影响。

从美国农业部 (USDA) 网站和美国国家海洋和大气管理局 (NOAA) 网站收集的数据，为 1991 年至 2006 年间堪萨斯州的大豆产量及相关环境变量提供了宝贵的信息。这些数据涵盖了一个重要的农业时期和地理区域，因为堪萨斯州是大豆生产的领先州之一。

在以下分析中，我们考虑三个预测变量，用于研究堪萨斯州各县的年度大豆产量。函数型预测变量记为 $X(t)$，表示日平均气温。我们具体考察 2 月至 11 月的温度数据，以探究寒冷月份和温暖月份的温度对大豆产量的影响。两个标量预测变量——年降水量和灌溉面积与收获土地面积之比 (Z2)——分别记为 Z_1 和 Z_2，提供了可能影响大豆生长的其他环境信息。最后，响应变量记为 Y，表示堪萨斯州各县以蒲式耳计的年度大豆产量。接下来，我们要研究的模型如下：

$$Q_Y(u|X(t), Z_1, Z_2) = \alpha_0(u) + \alpha_1(u)Z_1 + \alpha_2(u)Z_2 + \int_{\mathcal{T}} \beta(t,u)X(t)dt. \tag{10}$$

我们使用均匀分布于 5% 至 95% 之间的 19 个分位数拟合模型。图 1 ˆ 展示了使用所提出的同时估计方法得到的估计斜率函数 $\beta(t,\ u)$。

ˆ 在图1中，估计的效应函数 $\beta(t,u)$揭示了堪萨斯州日平均气温与大豆产量之间在各分位数上截然不同的时变关联。6月之前，温度似乎对所有分位数的产量均无实质性影响，表明生长季早期的温度总体上适宜大豆萌

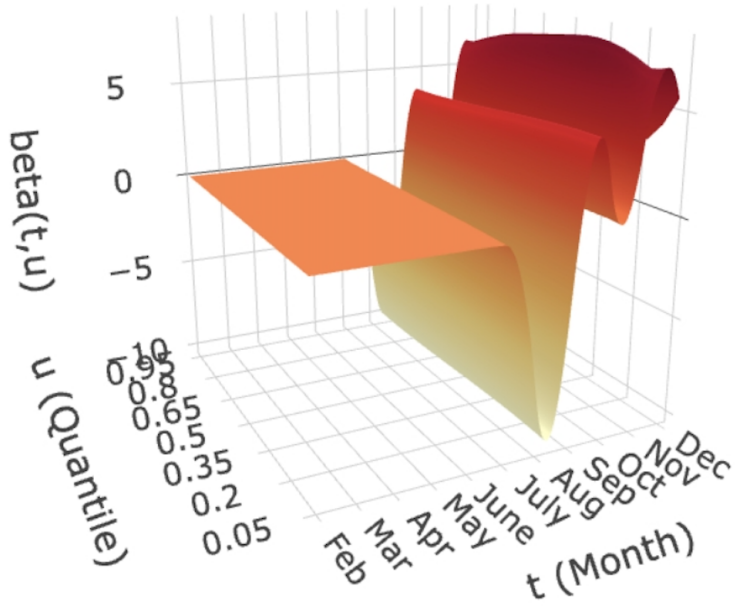

图1：基于堪萨斯州大豆产量数据集的同步函数分位数回归模型（10）的估计斜率函数 $\beta(t,u)$。

ˆ 发与初期生长。从7月至9月， $\beta(t,u)$在5%至95%分位数上持续为负，表明夏末较低的温度有利于大豆生长——这很可能是因为在豆荚灌浆等关键生长阶段减轻了热胁迫。从9月下旬至10月中旬， $\beta(t,u)$在所有分位数上转为正值，意味着初秋这一时期较高的温度可能促进持续生长与成熟。然而，从10月中旬开始，斜率函数下降，并在月底变为负值，表明此时反常温暖的条件可能阻碍生理成熟或干燥过程。有趣的是，从11月起， $\beta(t,u)$再次上升并保持为正，表明11月较暖的温度与较高产量相关。一种合理的解释是，在11月较温暖的年份，大豆作物——尤其是晚播品种或复种系统（例如，小麦后种植的大豆）——更不容易遭受早霜冻害，从而能够更充分地成熟并提高种子品质。这些发现凸显了温度的重要性

大豆产量的时机，并强调生长季后期发育对气候变异的敏感性。

ˆ 总体而言，估计量 $\beta(t,u)$ 为深入了解堪萨斯州日平均气温与大豆产量之间的动态关系提供了宝贵见解。具体而言，我们的研究结果表明，六月之前堪萨斯州的温度最适宜大豆萌发和生长。夏末（七月至九月）是一个关键时期，此期间较低的温度有利于大豆生长。秋季和初冬是另一个重要时期，此期间较高的温度有利于大豆生长。这些发现有助于制定策略，以优化目标地区的大豆生产。

5 结论与讨论

本文介绍了一种新方法，用于在函数型分位数回归模型 (1) 下估计二元斜率函数$\beta(t,u)$。其目标是研究并理解函数型预测变量与标量响应变量之间的动态关系。我们提出的方法采用局部稀疏估计技术，使我们能够识别出函数型预测变量对响应变量影响显著的特定定义域。

我们通过模拟比较了基于所提出方法和传统方法的函数型分位数回归模型中斜率函数估计的表现。数值结果表明，所提出的方法具有优越性。此外，我们还使用大豆产量数据集说明了所提出的方法与传统方法之间的差异。真实数据分析结果表明，所提出的方法能够提高模型可解释性。

与另一种常用的局部稀疏估计方法 fSCAD 相比，本文提出的方法的主要优势在于计算简单。与组

套索回归不同，fSCAD 正则化是非凸的，其数值优化通常依赖局部线性或局部二次近似，通过迭代使用凸替代函数近似非凸目标函数。这些迭代策略需要谨慎设定初始值和步长。初始点选择不当可能导致算法无法收敛，而步长设置不合适则会造成效率低下或解不理想：步长过小可能导致收敛缓慢，步长过大则可能使算法收敛到局部最优解，而非全局最优解。相比之下，组 LASSO 惩罚项不存在这些问题，因为它会形成凸优化问题，求解过程更加稳定且计算效率更高。

补充材料

FQRSupplementary.pdf: 包含模拟研究及手稿中理论结果证明的补充文件。

数据可用性声明： 支撑本研究结论的数据可在 `https://anonymous.4open.science/r/LocalFQR-1322/`公开获取。

参考文献

ApS, M. (2019). MATLAB 的 MOSEK 优化工具箱。用户指南和参考手册，版本，4(1):116。

Azzimonti, L., Sangalli, L. M., Secchi, P., Domanin, M., 和 Nobile, F. (2015). 通过带 PDE 惩罚的空间回归进行血流速度场估计。《美国统计学会杂志》，110:1057–1071。

Beyaztas，U.、Tez 和 Shang，H. L.（2024年）。稳健的标量对函数部分分位数回归。应用统计学杂志，51：1359–1377。

Cardot，H.、Crambes 和 Sarda，P.（2005年）。协变量为函数时的分位数回归。《非参数统计学杂志》，17(7)：841–856。

Cardot，H.、Ferraty，F. 和 Sarda，P.（2003年）。函数型线性模型的样条估计量。《中华统计学志》，13(3):571–591。

Chen，K. 和 M¨uller，H.-G.（2012年）。协变量为函数时的条件分位数分析及其在生长数据中的应用。《皇家统计学会志：*B*辑（统计方法学）》，74(1):67–89。

Fan，J. 和 Li，R.（2001年）。通过非凹惩罚似然进行变量选择及其预言性质。《美国统计学会杂志》，96(456):1348–1360。

Fang，K.、Zhang，X.、Ma, S. 和 Zhang, Q.（2020年）。多输出函数型线性回归的平滑与局部稀疏估计。《统计计算与模拟杂志》，90(2):341–354。

Fu，A.、Narasimhan，B. 和 Boyd，S.（2020年）。CVXR：用于规范凸优化的 R软件包。《统计软件杂志》，94:1–34。

Gurer，S.，Shang，H. L.，Mandal，A. 和 Beyaztas，U.（2024年）。标量对函数回归中的局部稀疏稳健偏最小二乘。《统计学与计算》，34(5)：150。

He，Q.，Kong, L.，Wang，Y.，Wang，S.，Chan，T. A. 和 Holland，E.（2016年）。异质稀疏性下的正则化分位数回归及其在数量遗传性状中的应用。《计算统计与数据分析》，95：222–239。

Hu，B.，Hu，X.，Liu，H.，You，J. 和 Cao, J.（2024年）。同步函数型分位数回归。《中华统计学志》，34(2)。

James，G. M.、Wang，J. 和 Zhu，J.（2009年）。可解释的函数型线性回归。《统计年鉴》，37(5A)：2083–2108。

Kato, K.（2012年）。函数型线性分位数回归中的估计。《统计年鉴》，40(6)：3108–3136。

Koenker, R. 和 Bassett Jr, G.（1978年）。回归分位数。《计量经济学》：计量经济学会期刊，第 33–50 页。

Lai, M.-J. 和 Schumaker, L. L.（2007年）。三角剖分上的样条函数。剑桥大学出版社，剑桥。

Lai, M. J. 和 Wang, L.（2013）。用于回归的双变量惩罚样条。《中华统计学志》，23:1399–1417。

Li, X.、Wang, L. 和 Wang, H.（2021年）。超高维图像对标量回归的稀疏学习与结构识别。《美国统计学会杂志》，116:1994–2008。

Li, Y.、Wang, F.、Wu, M. 和 Ma, S.（2022年）。用于全基因组关联研究及多性状分析的整合型函数线性模型。《生物统计学》，23(2):574–590。

Liang, W., Zhang, Q., 和 Ma, S.（2023年）。部分函数型交互模型的局部稀疏分位数估计。《计算统计与数据分析》，186:107782。

Lin, Z., Cao, J., Wang, L., 和 Wang, H.（2017年）。函数型线性回归模型的局部稀疏估计量。《计算与图形统计学杂志》，26(2):306– 318。

Liu, Y. 和 Wu, Y.（2011年）。使用核约束进行同步多重非交叉分位数回归估计。《非参数统计学杂志》，23(2):415–437。

Lu, Y., Du, J., and Sun, Z.（2014年）。函数型部分线性分位数回归模型。《计量学》，77(2):317–332。

Ma, H., Li, T., Zhu, H., and Zhu, Z.（2019年）。超高维函数型部分线性模型的分位数回归。《计算统计与数据分析》，129:135–147。

Neocleous, T. and Portnoy, S.（2008）。关于回归分位数函数的单调性。《统计与概率快报》，78(10):1226–1229。

Qingguo, T. and Kong, L.（2017年）。函数型线性半参数模型中的分位数回归。《统计学》，51(6):1342–1358。

Ramsay, J. O. and Silverman, B. W.（2005年）。函数型数据分析。Springer，纽约，第二版。

Reiss, P. T., Goldsmith, J., Shang，H. L., and Ogden, R. T.（2017年）。标量对函数回归的方法。《国际统计评论》，85:228–249。

Sangalli, L. M., Ramsay, J.O., and Ramsay, T. O.（2013）。空间样条回归模型。《皇家统计学会杂志》*B*辑，统计方法论，75:681–703。

Schlenker, W. 和 Roberts, M. J.（2009年）。非线性温度效应表明，气候变化将对美国作物产量造成严重损害。《美国国家科学院院刊》，106(37):15594–15598。

Schnabel, S. K. 和 Eilers, P. H.（2013）。使用分位数片同时估计分位数曲线。《*AStA*统计分析进展》，97(1):77–87。

Stone, C. J.、Hansen, M.、Kooperberg, C. 和 Truong, Y. K.（1997）。扩展线性建模中的多项式样条及其张量积（附讨论）。《统计年鉴》，25(4):1371–1470。

Takeuchi, I., Le, Q., Sears, T., Smola, A., 等（2006）。非参数分位数估计。

Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N., 和 Frieler, K.（2019年）。气候极端事件对全球农业产量的影响。《环境研究快报》，14(5):054010。

Wood, S. N. (2017)。广义加性模型：*R* 语言入门。Chapman and Hall/CRC。

Yao, F., Sue-Chee, S., 和 Wang, F.（2017年）。正则化部分函数型分位数回归。《多元分析杂志》，156：39–56。

Yu, D., Kong, L., 和 Mizera, I.（2016年）。用于神经影像数据分析的部分函数型线性分位数回归。《神经计算》，195：74–87。

Yu, D., Zhang, L., Mizera, I., Jiang, B., 和 Kong, L.（2019年）。含多个函数型预测变量的分位数回归中的稀疏小波估计。《计算统计与数据分析》，136：12–29。

Zhou, J., Wang, N.-Y., 和 Wang, N.（2013）。在子区域上系数函数取零值的函数型线性模型。《中华统计学志》，23(1)：25–50。

Zou, H. 和 Yuan, M. (2008). 复合分位数回归与预言模型选择理论。《统计年鉴》，36(3):1108–1126。