位置、尺度和形状的非齐次马尔可夫切换广义加性模型

Katharina Ammann[1*]，Timo Adam[1]，Jan-Ole Koslik[1]

[1]德国比勒费尔德大学

arXiv:2601.03760v1 [stat.ME] 7 Jan 2026

摘要

我们提出对位置、尺度和形状马尔可夫切换广义加性模型（MS-GAMLSS）的扩展，使协变量不仅能够影响状态依赖分布的参数，还能影响状态转移概率。传统的 MS-GAMLSS 将分布回归与隐马尔可夫模型相结合，通常假设转移概率具有时间齐次性（即为常数），因此无法使状态转换响应由协变量驱动的变化。我们通过将转移概率建模为协变量的平滑函数，克服了这一局限，从而能够灵活地、基于数据刻画协变量依赖的状态动态。估计在惩罚似然框架下进行，其中自动平滑度选择可控制模型复杂度并防止过拟合。我们通过模拟以及对汉莎航空每日股票价格和西班牙能源价格的应用，评估了所提出的方法。结果表明，将宏观经济指标纳入转移概率能够为市场动态提供更多见解。用于复现结果的数据和 R 代码可在线获取。

关键词：分布回归模型；位置、尺度和形状广义加性模型；隐马尔可夫模型；时间序列建模。

1 引言

近年来，位置、尺度和形状马尔可夫切换广义加性模型（MS-GAMLSS）已成为一类灵活的时间序列回归模型，它通过潜在状态过程捕捉条件响应分布的动态结构变化（Adam 等，2022）。传统上，控制该状态过程的转移概率被假定为时间齐次，即随时间保持不变。尽管这一假设简化了估计，但可能无法充分反映现实应用。例如，在金融市场中，平静和波动状态之间的转换往往由宏观经济状况驱动，而非纯粹随机发生（Foroni 等，2024）。本文提出一种对时间齐次 MS-GAMLSS 的扩展，其中潜在状态之间的转移概率取决于协变量。这一模型形式使我们能够描述潜在状态动态受到可观测条件系统性影响的情形。

*通讯作者；电子邮箱：katharina.ammann@uni-bielefeld.de。

所提出的方法建立在丰富的研究文献之上。自 Rigby 和 Stasinopoulos 于2005年提出 GAMLSS 以来，这类模型通过允许响应分布的参数——不仅是均值，还包括方差、偏度和峰度等——借助平滑函数依赖于协变量，革新了回归建模。其应用涵盖众多领域，包括洪水风险建模（Villarini 等人，2009年）、临床数据中的百分位数估计（Hossain 等人，2016年）、语音可懂度测试（Hu 等人，2015年；Stasinopoulos 等人，2024年），以及根据超声测量指标预测出生体重（Stasinopoulos 等人，2024年）。其他应用还包括电子商务需求预测（Ulrich 等人，2021年）、违约风险暴露建模（Wattanawongwan 等人，2023年）、分布不确定性下的资产定价（Regis 等人，2023年），以及选举结果建模（Balzer 等人，2025）。

GAMLSS 的灵活性也引起了各大机构的关注。例如，国际货币基金组织将 GAMLSS 纳入其针对美国经济的压力测试框架，由此深入了解系统性金融风险，并与美国财政部、美联储及其他主要机构合作，为政策建议提供依据（国际货币基金组织，2015年）。同样，英格兰银行运用 GAMLSS 评估英国抵押贷款市场的脆弱性，尤其关注尾部风险以及借款人特征的分布变化（英格兰银行，2017年）。在私营部门，美国银行等机构利用 GAMLSS 对内部信用风险模型进行基准评估，尤其是违约损失率和违约风险暴露等组成部分；对于这些部分，灵活地建模偏斜且厚尾的损失分布至关重要（Tong，2020年）。

尽管 GAMLSS 的大多数应用都假设观测值相互独立，但在时间序列数据中，这一假设往往不成立，因为时间依赖和状态转换行为十分常见（Goldfeld 和 Quandt，1973年；Hamilton，1989年）。为解决这一问题，Langrock、Adam 等人（2018年）提出了 MS-GAMLSS，这是一类将 GAMLSS 与隐马尔可夫状态结构相结合的时间序列回归模型。MS-GAMLSS 将 GAMLSS 框架扩展至时间序列，通过允许响应分布的参数随协变量变化，并由依赖于潜在不可观测状态的平滑函数加以描述。换言之，假设时间序列由若干独立的分布回归模型生成，每个模型对应一种不同状态。相关研究还包括 Hambuckers 等人（2018年）将复合泊松回归纳入马尔可夫切换模型，以及 Adam 等人（2022年）提出在 MS-GAMLSS 中采用梯度提升，以适应高维预测变量情形。在这条研究脉络的基础上，本文进一步扩展 MS-GAMLSS，明确将隐状态过程建模为一个依赖协变量、因而在时间上非齐次的马尔可夫链，其中协变量直接影响转移概率。

本文其余部分安排如下。第2节介绍所提出的方法，并说明该方法如何建模状态特定的分布动态和协变量依赖的状态转移。第3节通过模拟研究评估模型的性能。第4节给出两个案例研究，分别涉及汉莎航空股票的每日收盘价和西班牙的能源价格。可在 GitHub[1] 上获取数据和 `R` 用于复现结果的代码。

[1]https://github.com/ammann99/nonhomogeneousMSGAMLSS.

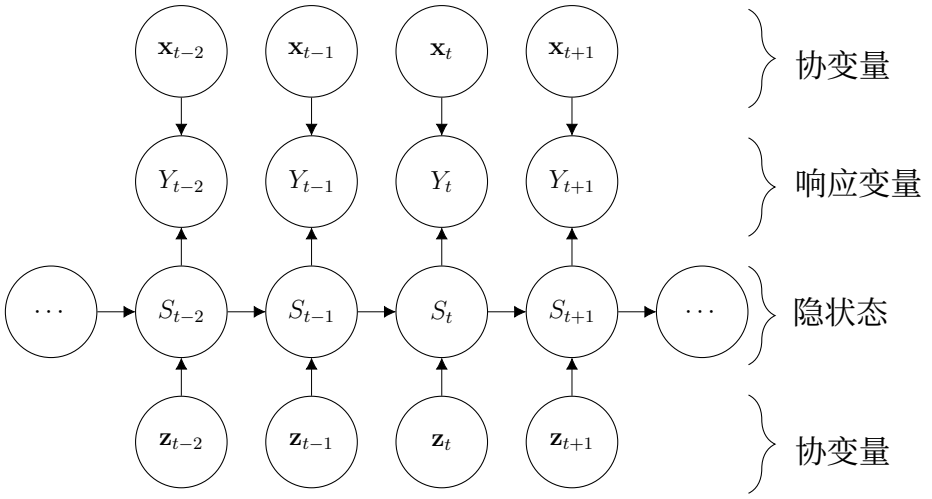

图 1：非齐次 MS-GAMLSS 依赖结构示意图。协变量同时影响隐状态过程 $\{S_t\}_{t=1,\ldots,T}$ 和观测到的状态依赖过程 $\{Y_t\}_{t=1,\ldots,T}$。

2 方法

2.1 模型表述与依赖结构

MS-GAMLSS 将表示潜在状态的隐马尔可夫链与可观测的状态依赖响应过程相结合（依赖结构示意见图1）。本小节介绍所提出的非齐次 MS-GAMLSS 的结构和协变量依赖的状态转移。

隐状态过程由离散时间的 N 状态马尔可夫链$\{S_t\}_{t=1,\ldots,T}$ 表示。假设满足一阶马尔可夫性，即

$$\Pr(S_{t+1} \mid S_1, \ldots, S_t) = \Pr(S_{t+1} \mid S_t), \quad t = 1, \ldots, T-1.$$

这简化了依赖结构，并实现了高效的似然计算；如有必要，也可以进行高阶扩展（Zucchini 等人，2016年）。

初始状态概率表示在时间 $t = 1$ 开始时处于各个状态的概率，并汇总于向量 $\boldsymbol{\delta}^{(1)} = (\delta_1, \ldots, \delta_N)$ 中，其中 $\delta_i = \Pr(S_1 = i)$。对于齐次马尔可夫链，状态转移由时不变的$N \times N$ 转移概率矩阵（t.p.m） $\boldsymbol{\Gamma} = (\gamma_{i,j})_{i,j=1,\ldots,N}$ 表示，其中 $\gamma_{i,j} = \Pr(S_{t+1} = j \mid S_t = i)$,，对于 $i, j = 1, \ldots, N$。

为使协变量能够影响状态转换动态，我们考虑非齐次（即随时间变化的）转移概率。这些概率由依赖时间的 t.p.m. $\boldsymbol{\Gamma}_t = (\gamma_{i,\ j}^{(t)})_{i,\ j=1.....,\ N}$ 描述，其中

$$\gamma_{i,j}^{(t)} = \Pr(S_{t+1} = j \mid S_t = i, \mathbf{z}_t), \text{ for } i, j = 1, \ldots, N,$$

其中，$\mathbf{z}_t = (z_{1t}, \ldots..., \ z_{Kt})^\top$ 表示影响状态转移的协变量。

A mu采用多项逻辑链接函数将加性预测子与转移概率联系起来，即 -

转移概率，即

$$\gamma_{i,j}^{(t)} = \frac{\exp\big(\eta_{i,j}^{(t)}(\mathbf{z}_t)\big)}{\sum_{l=1}^{N}\exp\big(\eta_{i,l}^{(t)}(\mathbf{z}_t)\big)}, \quad \text{for } i,j = 1,\ldots,N,$$

其中，对于 $i \neq j$，线性预测子具有以下形式

$$\eta_{i,j}^{(t)}(\mathbf{z}_t) = \beta_0^{(i,j)} + \sum_{k=1}^{K}\beta_k^{(i,j)} z_{kt} + \sum_{k=1}^{K} s_k^{(i,j)}(z_{kt}),$$

且 $\eta_{i\,i}^{(t)}{}_{,}(\mathbf{z}_t) = 0$ 以确保可识别性。$\beta_0^{(i,j)}$ 是状态对特异截距，$\beta_k^{(i,j)}$ 是线性系数，而 $s_k^{(i,j)}(\cdot)$ 是使用样条基函数建模的平滑函数（见公式（2））。

假设观测到的状态依赖过程$\{Y_t\}_{t=1,\ldots,T}$在给定状态的条件下相互独立，并服从某个参数分布 $\mathcal{D}$,

$$Y_t \sim \mathcal{D}\big(\mu_t^{(s_t)}, \sigma_t^{(s_t)}, \nu_t^{(s_t)}, \tau_t^{(s_t)}\big), \tag{1}$$

其中 $\mu_t^{(s_t)}$、$\sigma^{(s_t)}{}_t$、$\nu^{(s_t)}{}_t$ 和 $\tau^{(s_t)}{}_t$ 分别是状态特异、协变量依赖参数，表示均值、方差、偏度和峰度。这些参数通过单调链接函数 $g_k(\cdot)$ 与加性预测子相连接，使得

$$\begin{aligned}
g_1\big(\mu_t^{(i)}\big) &= \mathbf{X}_1\boldsymbol{\beta}_1^{(i)} + \sum_{j=1}^{J_1} s_{1,j}^{(i)}(x_{j,t}),\\
g_2\big(\sigma_t^{(i)}\big) &= \mathbf{X}_2\boldsymbol{\beta}_2^{(i)} + \sum_{j=1}^{J_2} s_{2,j}^{(i)}(x_{j,t}),\\
g_3\big(\nu_t^{(i)}\big) &= \mathbf{X}_3\boldsymbol{\beta}_3^{(i)} + \sum_{j=1}^{J_3} s_{3,j}^{(i)}(x_{j,t}),\\
g_4\big(\tau_t^{(i)}\big) &= \mathbf{X}_4\boldsymbol{\beta}_4^{(i)} + \sum_{j=1}^{J_4} s_{4,j}^{(i)}(x_{j,t}),
\end{aligned}$$

其中 $\mathbf{X}_k$ 表示线性效应的设计矩阵，$\boldsymbol{\beta}_k^{(i)}$ 表示状态特异的线性回归系数，而 $s_k^{(i)}{}_{,\,j}(\cdot)$ 表示平滑协变量效应。每个加性预测子中的协变量或平滑项数量记为 J_k。每个分布参数都通过单调链接函数 $g_k(\cdot)$ 进行变换，映射到实值加性预测子上，从而确保满足参数约束（例如，方差为正）。

状态和状态依赖过程中的所有平滑效应均表示为 M 个固定样条基函数的有限线性组合，即:

$$s(x) = \sum_{m=1}^{M} b_m\phi_m(x), \tag{2}$$

采用样条系数 b_m 和样条基函数 $\phi_m(x)$（Boor，1978；Eilers 和 Marx，1996）。

这种灵活的模型结构支持针对时间序列数据进行状态依赖且由协变量驱动的分布回归。尽管给定状态后观测值的条件独立假设可能显得过于严格，但状态过程会显式地引入序列依赖这一事实为该假设提供了支持。

2.2 似然计算与模型拟合

由于似然涉及对所有可能的状态序列求和，直接计算很快就会变得不可行。对于马尔可夫切换模型，可以方便地使用前向算法来计算似然；该算法利用所谓的 前向变量 递归计算似然（Zucchini 等人，2016年）。前向变量 $\alpha_t(j)$ 表示截至时间 t 的观测值以及在时间 t 处于状态 j 的联合密度，定义为 $\alpha_t(j) = f(y_1, \ldots, y_t, S_t = j)$。前向变量可组合成向量 $\boldsymbol{\alpha}_t = (\alpha_t(1)\ \alpha_t(2) \ldots, \alpha_t(N))$。

前向算法的关键思想是递归更新这些前向变量，从而能够在线性时间内进行似然计算。在 $t = 1$ 时，前向变量初始化为

$$\boldsymbol{\alpha}_1 = \boldsymbol{\delta}^{(1)}\mathbf{P}(y_1),$$

其中 $\mathbf{P}(y_t) = \text{diag}(f(y_t \mid \mu_t^{(1)}, \sigma_t^{(1)}, \nu_t^{(1)}, \tau_{(1)t}) \ldots, f(y_t \mid \mu_t^{(N)}, \sigma_t^{(N)}, \nu_t^{(N)}, \tau_{(N)t}))$ 是一个包含状态依赖响应密度的对角矩阵。

对于后续时间步 $t = 2$、.....、T，通过递推关系迭代更新前向变量

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1}\boldsymbol{\Gamma}^{(t)}\mathbf{P}(y_t).$$

根据全概率公式，似然等于时间 T 的前向概率之和

$$\mathcal{L}(\boldsymbol{\theta}) = f(y_1, \ldots, y_T) = \sum_{i=1}^{N} \alpha_T(i) = \boldsymbol{\alpha}_T\mathbf{1}^\top,$$

其中，$\boldsymbol{\alpha}_T$ 是时间 T 的前向变量向量，$\mathbf{1}^\top$ 是由 1 组成的行向量。

为解决长时间序列的似然计算过程中出现的数值问题（如下溢或上溢），我们采用缩放策略计算对数似然，并使用经过标准化以避免数值不稳定性的标准化前向变量（Lystig 和 Hughes，2002年；有关缩放前向算法的详细信息，请参见附录附录 A）。

为防止过拟合，我们通过惩罚相关系数来限制方程 (2) 中平滑函数的灵活性。设 $\boldsymbol{b}_p$ 表示单个平滑函数的样条系数向量。为简化记号，我们使用单一索引。假设状态依赖分布参数和转移概率矩阵（t.p.m.）总共包含 P 个平滑函数。

然后，可以通过最大化

$$\log \mathcal{L}(\boldsymbol{\theta}) - \frac{1}{2}\sum_{p=1}^{P} \lambda_p \boldsymbol{b}_p^\top \boldsymbol{S}_p \boldsymbol{b}_p, \tag{3}$$

来拟合模型，其中 log $\mathcal{L}(\boldsymbol{\theta})$ 使用缩放前向算法计算，λ_1、……、λ_P为平滑参数，$\boldsymbol{S}_1$、……、$\boldsymbol{S}_P$为每个平滑函数对应的惩罚矩阵。对于简单的P 样条 惩罚（Eilers 和 Marx，1996），每个 $\boldsymbol{S}_p$ 都可以取为 $\boldsymbol{D}^\top\boldsymbol{D}$，其中 $\boldsymbol{D}$ 为二阶差分矩阵。Wood（2017）详细介绍了一种构造这些矩阵的更一般方法。显然，每个 λ_p 都决定了第 p 个平滑函数的平滑度。对于 $\lambda_p = 0$，该函数不受约束；而当 $\lambda_p \to \infty$ 时，该函数则趋于线性函数。实际应用中，方程（2）中的基函数取值和惩罚矩阵 $\boldsymbol{S}_p$ 都可以方便地通过基于模型公式的 `mgcv` `R` 包（Wood，2017）获得，从而在基函数的具体选择上提供了很大的灵活性。

此处采用近似限制最大似然进行自动化、数据驱动的平滑度选择。Koslik（2024年）首次将这一方法引入马尔可夫切换模型。其核心思想是，公式 (3) 中的惩罚项为每个 $\boldsymbol{b}_p$ 施加了一个非正规多元高斯分布，因此可以将其视为随机效应。具体而言，公式 (3) 表明 $\boldsymbol{b}_p \sim \mathcal{N}(0,\ \boldsymbol{S}_p^-/\lambda_j)$，其中 $\boldsymbol{S}_p^-$ 是 $\boldsymbol{S}_p$ 的 Moore-Penrose 伪逆。因此，可以根据限制似然函数估计 $\boldsymbol{\lambda} = (\lambda_1、……、\lambda_P)$，同时将 $\boldsymbol{b}_1$、……、$\boldsymbol{b}_P$及所有固定效应积分消去。直观来说，这个限制似然衡量的是基于从各自分布中抽取的许多组$\boldsymbol{b}_1$、……、$\boldsymbol{b}_P$所能达到的平均似然。只有平滑度适当时，这个限制似然才会较大。如果 λ_p 非常小，抽取出的平滑函数就会非常曲折，产生一些很高的似然值，但也会有许多很低的值。如果λ_p 非常大，大多数抽取出的函数就会非常平直，因此大多数似然值都会相当小。

在计算方面，难以处理的高维积分通过所谓的拉普拉斯近似进行近似；该方法基于在给定 $\boldsymbol{\lambda}$的情况下，对式（3）的众数进行二阶泰勒近似。最终，这将问题简化为根据后续的惩罚拟合，对平滑度参数进行迭代更新。通常需要进行适量的惩罚模型拟合（10–50 次），才能找到合适的平滑参数向量。与使用网格搜索的简单替代方法相比，这种方法的计算成本低得多，尤其是在需要同时估计多个函数的平滑度时。此外，迭代过程也有其优势，因为可以将平滑度参数初始化为较大的值，从而得到非常稳定的初始惩罚拟合。在优化过程中，这些参数通常会逐渐减小，从而提高灵活性；但对于每次新的惩罚拟合，$\boldsymbol{\theta}$ 都可以用上一次迭代得到的惩罚估计进行初始化。这一策略使该过程既高效，又能相对稳健地避免收敛到惩罚似然的局部最优解。更详细的说明请参见 Koslik（2024年）。

在实际实现中，我们仅使用前向算法和 `penalty` 函数，在 `R` 中指定惩罚负对数似然函数；该函数由 `R`包 `LaMa` 实现（Koslik，2025）。随后，我们使用同一软件包中的 `qreml()` 函数联合估计模型和平滑度参数。在内部，该函数利用 `R` 包 `RTMB`提供的自动微分机制（Kristensen，2025），基于 C++ tape 快速计算式（3）及其梯度。

3 模拟研究

为评估所提出方法的性能，我们进行了以下模拟研究。潜在状态序列$\{S_t\}_{t=1,\dots,T}$由一个初始状态分布为 $\boldsymbol{\delta}^{(1)} = (0.5, 0.5)$ 的二状态、时间非齐次马尔可夫链生成。转移概率取决于均匀分布的协变量 $z_{1,t} \sim \mathcal{U}(-1, 1)$，因此时刻 t的转移概率矩阵为

$$\mathbf{\Gamma}_t = \begin{pmatrix} \gamma_{1,1}^{(t)} & \gamma_{1,2}^{(t)} \\ \gamma_{2,1}^{(t)} & \gamma_{2,2}^{(t)} \end{pmatrix},$$

每一行均采用多项逻辑回归设定，并包含协变量的线性和二次效应：

$$\operatorname{logit}\big(\gamma_{i,j}^{(t)}\big) = \eta_{i,j}(z_{1,t}), \quad \text{where} \quad \begin{aligned} \eta_{1,2}(z_{1,t}) &= -1.8 + 1.5\, z_{1,t} - 2\, z_{1,t}^2, \\ \eta_{2,1}(z_{1,t}) &= -2.1 - 2\, z_{1,t} - 1\, z_{1,t}^2, \end{aligned}$$

以及 $\eta_{i\,i}^{(t)}{}_{,}(z_{1,\ t}) = 0$ 以确保可识别性。给定隐状态 S_t，观测值 Y_t 服从条件正态分布，其特定状态的位置和尺度参数建模为第二个协变量 $x_{1,\ t} \sim \mathcal{U}(-1,\ 1)$ 的函数，其中 $Y_t \mid \{S_t = i\} \sim \mathcal{N}(\mu_t^{(i)},\ (\sigma_t^{(i)})^2)$，条件均值和对数标准差的预测变量定义如下

$$\begin{aligned} \mu_t^{(1)} &= x_{1,t} + x_{1,t}^2, & \log\big(\sigma_t^{(1)}\big) &= -0.5 + x_{1,t}^2, \\ \mu_t^{(2)} &= -1 + x_{1,t} + 0.5\sin(\pi x_{1,t}), & \log\big(\sigma_t^{(2)}\big) &= \log(1 + 0.5x_{1,t}). \end{aligned}$$

在每次 $R = 200$次独立模拟运行中，我们生成了长度为$T = 4,\ 000$ 的时间序列。所有模型均按照第 2.2 节所述，通过惩罚最大似然进行拟合。

模拟结果如图2所示。条件均值的估计平滑函数 $\hat{}\mu_t^{(i)}$ 、对数标准差 $\log(\hat{}\sigma_t^{(i)})$ 以及转移概率 $\hat{}\gamma_{i\,i}^{(t)}{}_{,}$ 均能很好地还原数据生成过程的真实函数关系。对于位置和尺度参数，估计值的变异性较低，且在整个协变量域内均紧密贴合真实函数。尤其是，在两种状态下，依赖状态的均值 $\hat{}\mu_t^{(i)}$ 和对数标准差 $\log(\hat{}\sigma_t^{(i)})$ 函数均具有良好的可识别性；协变量边界附近的变异性略有增加，但总体上与真实值高度吻合。对于估计的转移概率，函数 $\hat{}\gamma_{1\,1}^{(t)}{}_{,}$ 和 $\hat{}\gamma_{2\,2}^{(t)}{}_{,}$ 能够很好地近似真实转移模式，仅在协变量域边界附近（数据较稀疏处）存在轻微偏差。估计得到的平滑函数捕捉到了转移机制中的协变量效应，表明所提出的方法能够还原依赖协变量的切换行为，以及底层的分布结构和潜在动态。

4 实证应用

本节通过两个案例研究，展示所提出的非齐次 MS-GAMLSS 框架的应用。第一个案例分析汉莎航空股票每日收盘价的动态变化。第二个案例使用西班牙能源价格数据，展示该建模方法在不同经济环境下的灵活性。

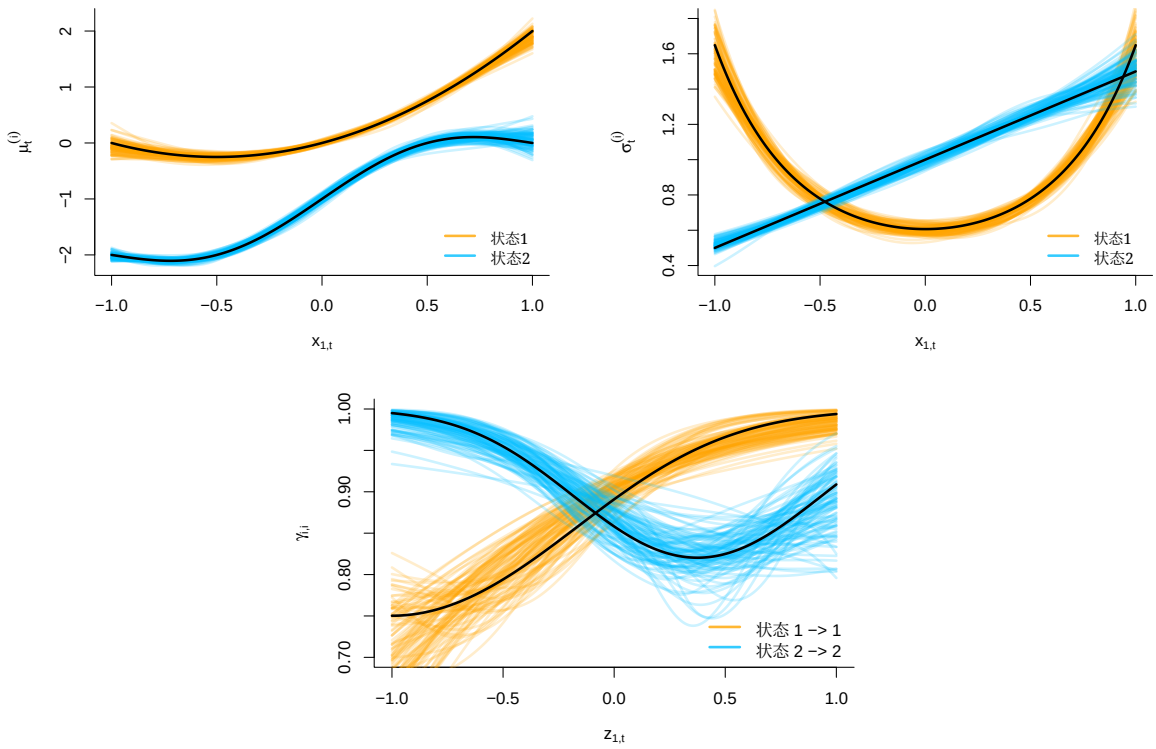

图2：估计的状态依赖效应和转移概率。图中展示了基于 200 次模拟运行得到的估计状态依赖函数：均值（左上图）、对数标准差（右上图）以及转移概率（下图）。状态 1 和状态 2 的估计结果分别以橙色和蓝色显示。真实函数以黑色实线表示。下图展示了估计的状态 1 和状态 2 的留存概率及其对应的真实曲线。

4.1 汉莎航空股票价格

本分析基于2007年8月1日至2025年6月27日期间汉莎航空股票价格的6,515个每日观测值。汉莎航空是欧洲最大的航空公司之一，也是德国DAX指数的成分股。样本期涵盖了多次重大金融危机，包括全球金融危机（2008年）、欧洲主权债务危机（2010年）、COVID-19疫情导致的经济下行（2020年），以及随后出现的市场波动性显著加剧的时期。这些事件为考察不同市场环境下的状态动态提供了丰富的背景。

协变量的选择依据的是航空公司股票表现、大宗商品价格与宏观经济状况之间已有充分文献记载的关系。布伦特原油价格以布伦特原油期货合约表示，作为唯一协变量纳入模型，同时影响收益的条件均值和方差，因为燃料成本是航空公司运营费用的重要组成部分。因此，石油价格的波动会对航空公司盈利能力产生直接且显著的影响（德国汉莎航空股份公司，2024年）。此外，期限利差（长期利率与短期利率之差）作为隐状态过程中的协变量，用以考虑由宏观经济状况驱动的状态转换动态。利差收窄或倒挂通常表明衰退风险上升，而利差扩大通常意味着人们预期经济增长（Estrella 和 Hardouvelis，1991）。

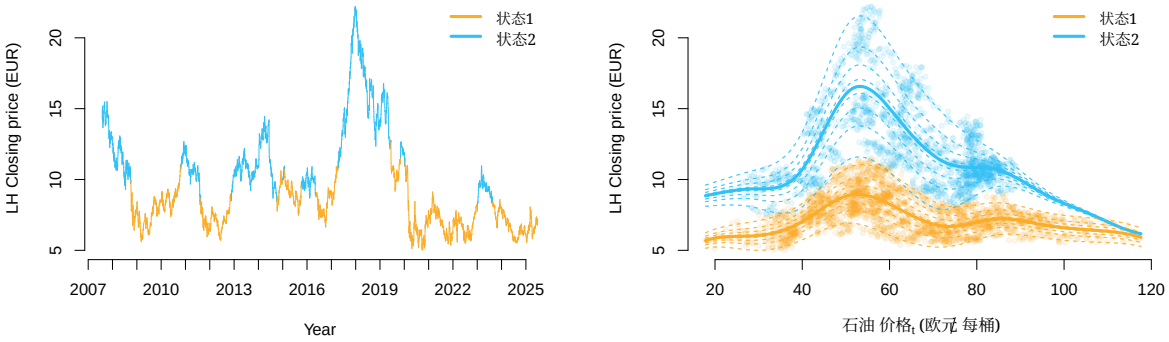

图3：收盘价的解码时间序列和估计的协变量效应。 左图：拟合的非齐次 MS-GAMLSS 下的解码时间序列。观测值按最可能的隐状态着色。右图：收盘价的状态依赖条件均值随原油价格变化的估计结果。实线表示各状态下估计的条件期望。虚线显示拟合的正态分布在一组规则概率水平（5%–95%）上的条件分位数。

在本应用中，期限利差采用滞后360天，以考虑宏观经济状况（如衰退）在收益率曲线发生变化后通常需要一段时间才会显现这一特点。

汉莎航空收盘价的模型设定如下：

$$\begin{aligned}
\text{LH Closing Price}_t \mid \{S_t = i\} &\sim \mathrm{N}\Big(\mu_t^{(i)},\, \sigma_t^{(i)}\Big)\,, \\
\mu_t^{(i)} &= \beta_{0,\mu}^{(i)} + s_{1,\mu}^{(i)}\big(\text{oil_price}_t\big), \\
\log\big(\sigma_t^{(i)}\big) &= \beta_{0,\sigma}^{(i)} + s_{1,\sigma}^{(i)}\big(\text{oil_price}_t\big), \\
\operatorname{logit}\big(\gamma_t^{(i,j)}\big) &= \beta_{0,\gamma}^{(i,j)} + s_{1,\gamma}^{(i,j)}\big(\text{spread}_{t-360}\big),
\end{aligned} \tag{4}$$

其中 $i, j = 1, 2$ 和 $i \neq j$。这里，$LH\ Closing\ Price_t$ 表示汉莎航空股票的每日收盘价，石油价格$_t$ 表示原油价格，利差$_{t-360}$ 表示一年前的期限利差。平滑函数 $s_{\cdot}^{(i)}(\cdot)$ 刻画协变量对状态相关均值和尺度的非线性效应，而转移概率则建模为期限利差的函数，并采用（状态特定的）logit 链接函数。

图3中的解码时间序列显示出两种不同状态。状态1（橙色）对应收盘价较低且价格走势更稳定的时期，而状态2（蓝色）则反映出较高的价格和更大的波动。状态序列是使用维特比算法（Viterbi，1967）得到的，该算法根据观测数据和估计的模型参数，找出最可能的隐状态序列。

原油价格（石油价格$_t$）对条件参数的估计平滑效应如图4所示。在两种状态下，每日收盘价的条件均值最初都会随着石油价格上涨而大幅增加。总体而言，在整个石油价格范围内，状态2对应着更高的条件均值水平；当石油价格接近每桶60欧元时，收盘价在约16欧元处达到峰值。相比之下，状态1始终处于较低水平，平均在7欧元左右波动。对于条件标准

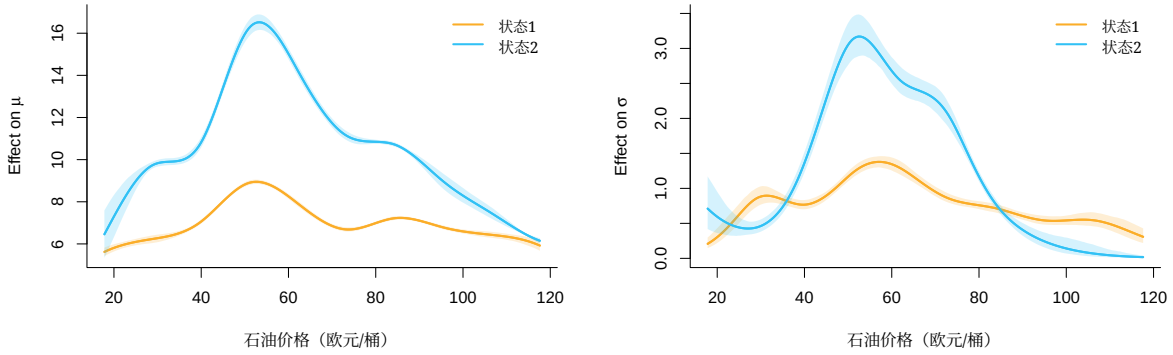

图4：石油价格的估计状态依赖效应。汉莎航空股票每日收盘价的条件均值（左）和标准差（右）所受的影响。线条表示估计的状态特定预测变量，阴影区域表示根据拟合的两状态正态MS-GAMLSS中状态依赖参数的后验分布得到的95%点态置信区间。

差，状态1中的数值主要保持在 σ 尺度的下半部分。相比之下，状态2对波动率的影响通常更强，尤其是在石油价格的中等区间，此时对 σ 的影响大约可以翻倍。

图5展示了估计的状态转移概率随滞后期限利差变化的情况。左图显示保持在当前状态的概率，右图则显示状态之间切换的概率。阴影区域表示平稳分布的95%点态置信区间，该区间根据转移模型参数的后验样本计算，参数后验样本基于观测到的费舍尔信息矩阵得出；对每个样本点计算平稳分布，并以经验分位数进行汇总。实线表示基于最大惩罚似然估计得到的估计转移概率。期限利差越高，状态1的持续性越低；期限利差越低，转入状态2的情况则越少见。

图5底部展示了在给定估计转移概率的情况下，平稳分布如何随利差变化。当利差为负时，处于状态1的概率较高，这与此前的结果一致，即状态1对应较低的价格水平。随着利差增加，该概率逐渐下降，直到利差略低于3时，两种状态的概率相交，状态2变得更有可能出现。这一模式符合这样的解释：较高的利差与正向经济预期相关。

较宽的置信区间反映出，解码状态序列（见图3）仅包含少数可能的状态转换，因此可用于估计与离散程度相关的转换行为的信息有限。

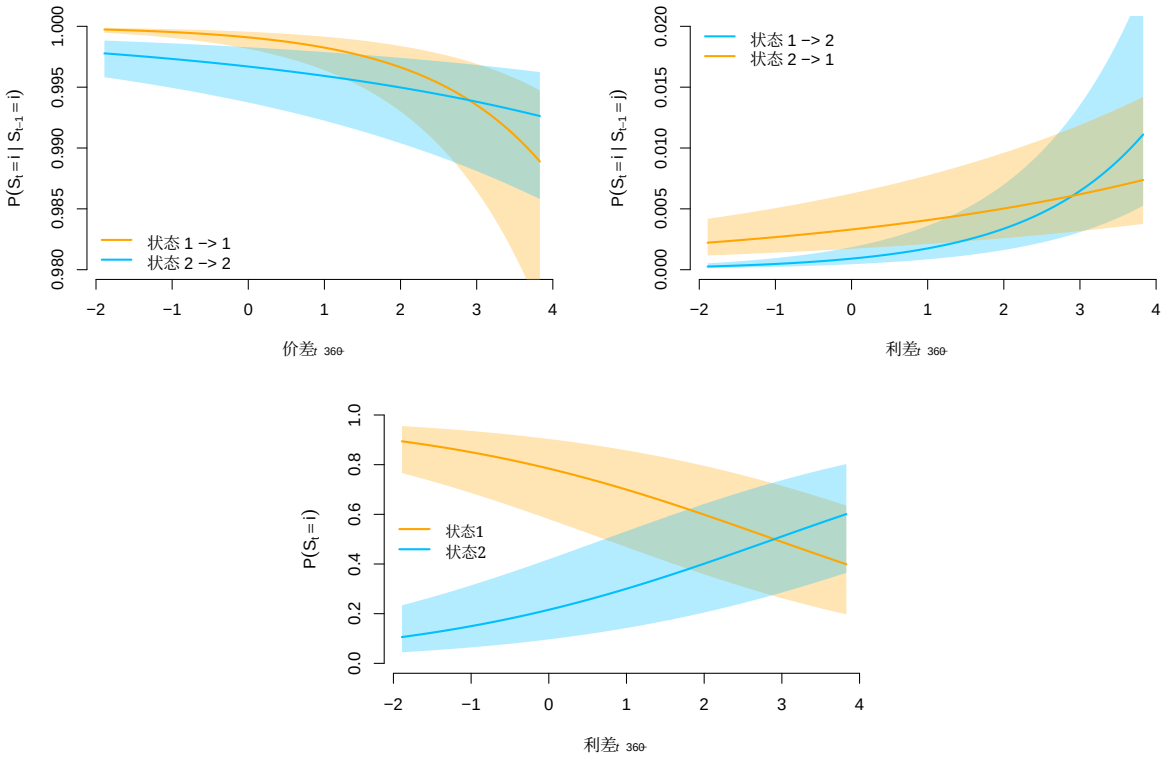

图5：估计的转移概率随滞后期限利差变化的函数。 左：保持在同一状态的概率。右：状态间切换的概率。底部：依赖协变量的平稳状态概率随滞后期限利差变化的函数。阴影区域表示根据依赖协变量的转移概率后验分布得出的95%点态置信区间。

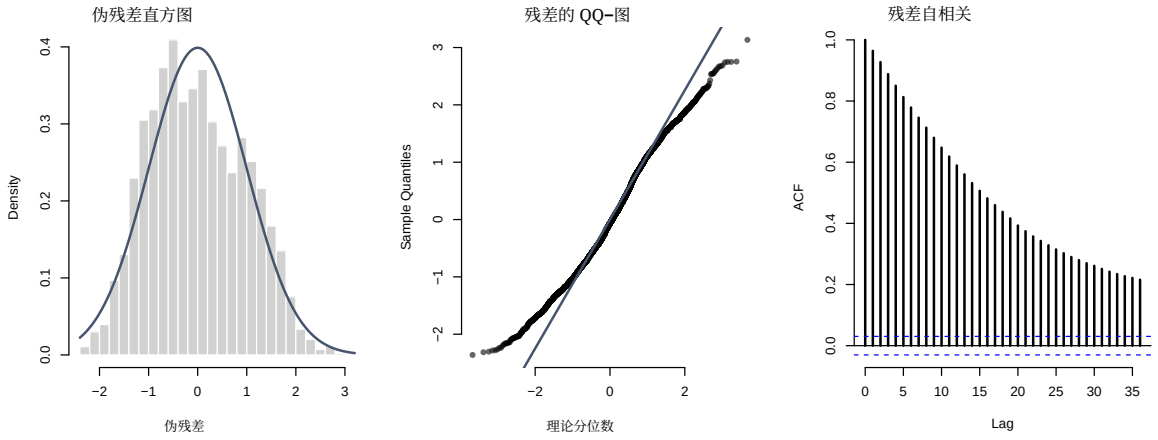

图6：汉莎航空拟合模型的残差诊断。 左：伪残差直方图，并叠加标准正态密度，表明其与理论分布总体上较为吻合。中：伪残差QQ图，突出显示极端分位数处的偏差。右：伪残差的自相关函数，显示仍存在残差自相关；鉴于模型中未显式包含自回归成分，这一情况是预期之中的。

为评估拟合模型的充分性，我们检查了伪残差诊断，如图6所示。伪残差通过概率积分构造

经过变换后，若模型设定正确，数据应服从标准正态分布。直方图表明，尾部整体拟合良好，但相对于标准正态分布，中心部分略有偏移。然而，QQ 图显示极端分位数存在明显偏差，说明正态状态依赖分布未能充分适应观测到的非对称性。

伪残差的自相关函数显示仍存在序列依赖。这一结果是预期之中的，因为模型未包含自回归成分或其他用于捕捉短期时间依赖的显式机制。因此，观测到的残差自相关并不表明条件分布设定有误。尽管加入自回归结构并不困难，但为保持模型设定简约，并将重点放在分布建模而非额外的动态复杂性上，此处省略了这一扩展。此模式促使我们在接下来的应用中采用稍具灵活性的状态依赖分布，并使用偏正态设定对偏度进行显式建模。

4.2 西班牙能源价格

作为第二个应用案例，我们分析西班牙的每日平均电价，以展示建模框架在不同经济环境中的应用。遵循此前针对该数据集的研究（Sanchez 等，2009；Langrock、Kneib 等，2017；Adam 等，2022；Koslik，2024），我们重点关注相关经济协变量驱动的条件分布动态和状态转换行为建模。数据可通过 R软件包 `MSwM`(Sanchez-Espigares 和 Lopez-Moreno，2021)获取，包含从2002年2月4日至2008年10月31日的1,760个工作日。

为保持模型结构简单且易于解释，我们将各状态下的条件均值、尺度和偏度建模为仅由石油价格决定的函数。此外，欧元兑美元汇率被纳入隐状态过程，使状态转换能够与汇率波动相关联。在这一设定中，石油价格反映全球大宗商品市场动态，而汇率变动则体现推动能源价格状态变化的更广泛宏观经济影响。

西班牙能源价格的二状态 MS-GAMLSS 设定如下：

$$
\begin{aligned}
\text{Energy Price}_t \mid \{S_t = i\} &\sim \text{SN}\Big(\mu_t^{(i)},\, \sigma_t^{(i)},\, \nu_t^{(i)}\Big), \\
\mu_t^{(i)} &= \beta_{0,\mu}^{(i)} + s_{1,\mu}^{(i)}(\text{oil_price}_t), \\
\log(\sigma_t^{(i)}) &= \beta_{0,\sigma}^{(i)} + s_{1,\sigma}^{(i)}(\text{oil_price}_t), \\
\nu_t^{(i)} &= \beta_{0,\nu}^{(i)} + s_{1,\nu}^{(i)}(\text{oil_price}_t), \\
\text{logit}\big(\gamma_{i,j}^{(t)}\big) &= \beta_{0,\gamma}^{(i,j)} + s_{1,\gamma}^{(i,j)}(\text{exchange_rate}_t),
\end{aligned}
\tag{5}
$$

使用 $i, j = 1, 2$ 和 $i \neq j$。 $Energy\ Price_t$表示西班牙的每日平均电价，石油价格$_t$ 表示原油价格，汇率$_t$ 表示美元与欧元之间的汇率。-在这里，我们假设 $Energy\ Price_t$ 服从偏态正态分布，并拟合了一个二状态 MS-GAMLSS，其中条件均值、条件方差和条件偏度采用状态依赖预测变量。平滑函数 $s^{(i)}(\cdot)$ 刻画非线性

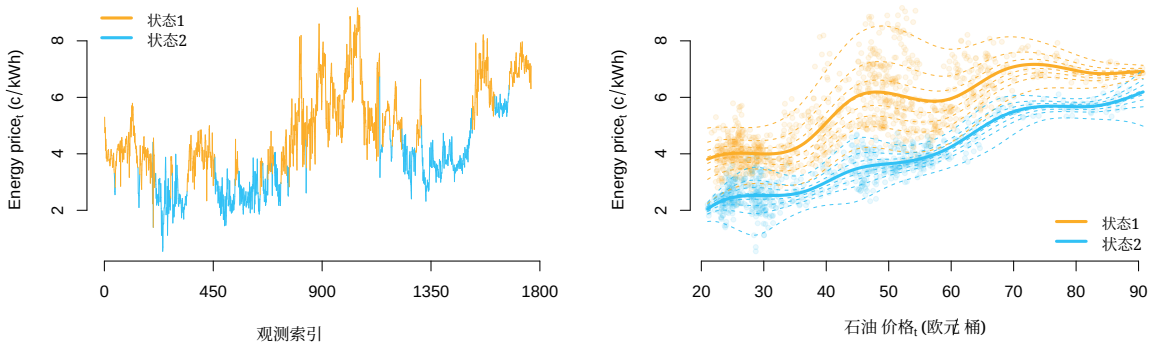

图7：能源价格解码时间序列及估计的协变量效应。左图：拟合的非齐次 MS-GAMLSS 下的解码时间序列，观测值按最可能的隐状态着色。右图：能源价格的状态依赖条件均值随石油价格变化的估计结果。实线表示各状态下估计的条件期望，虚线表示在规则概率水平网格（1%–99%）上拟合的偏态正态分布的条件分位数。

协变量对状态依赖参数的效应，而转移概率则通过 logit 链接函数建模为汇率的状态特定函数。

图7中的解码时间序列揭示了两种不同状态。状态1（橙色）对应电价较高且波动较大的时期，而状态2（蓝色）反映了价格较低且更加稳定的时期。

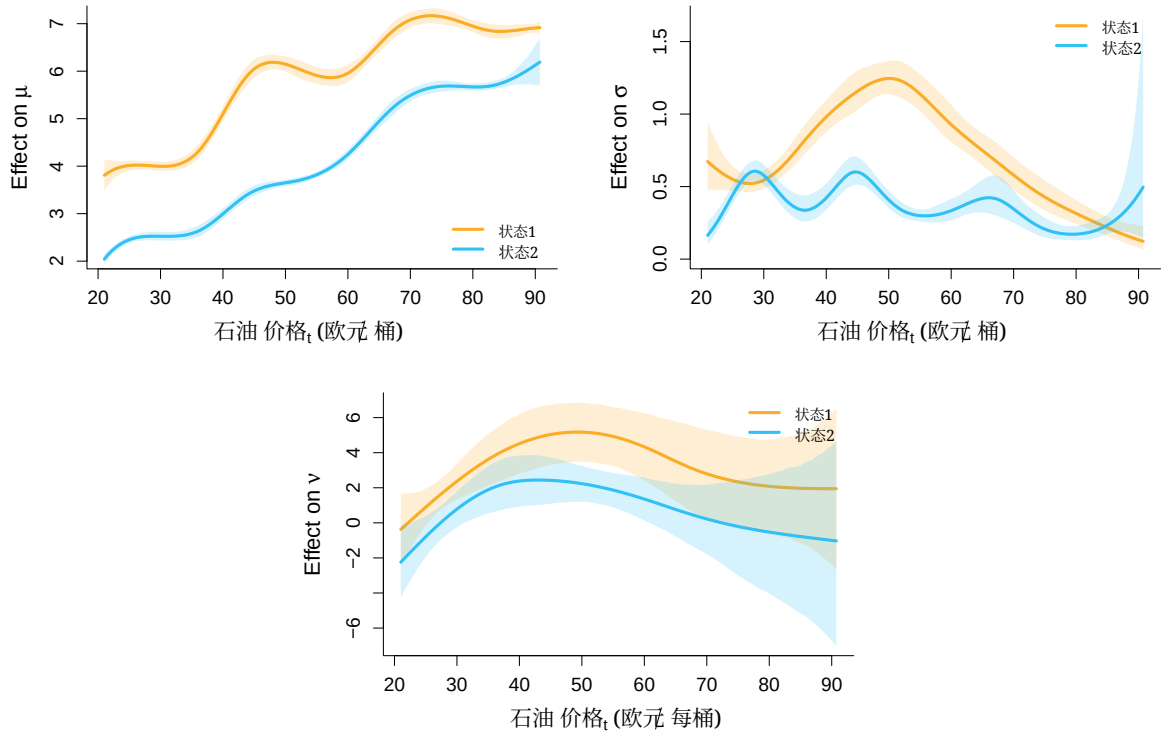

图8：原油价格的状态依赖平滑效应估计。左上：对每日能源价格条件均值的影响。右上：对条件尺度的影响。下方：对条件偏度的影响。曲线表示估计的状态特定预测变量；阴影区域表示根据拟合的双状态偏态正态 MS-GAMLSS 中状态依赖参数的后验分布得出的95%点态置信区间。

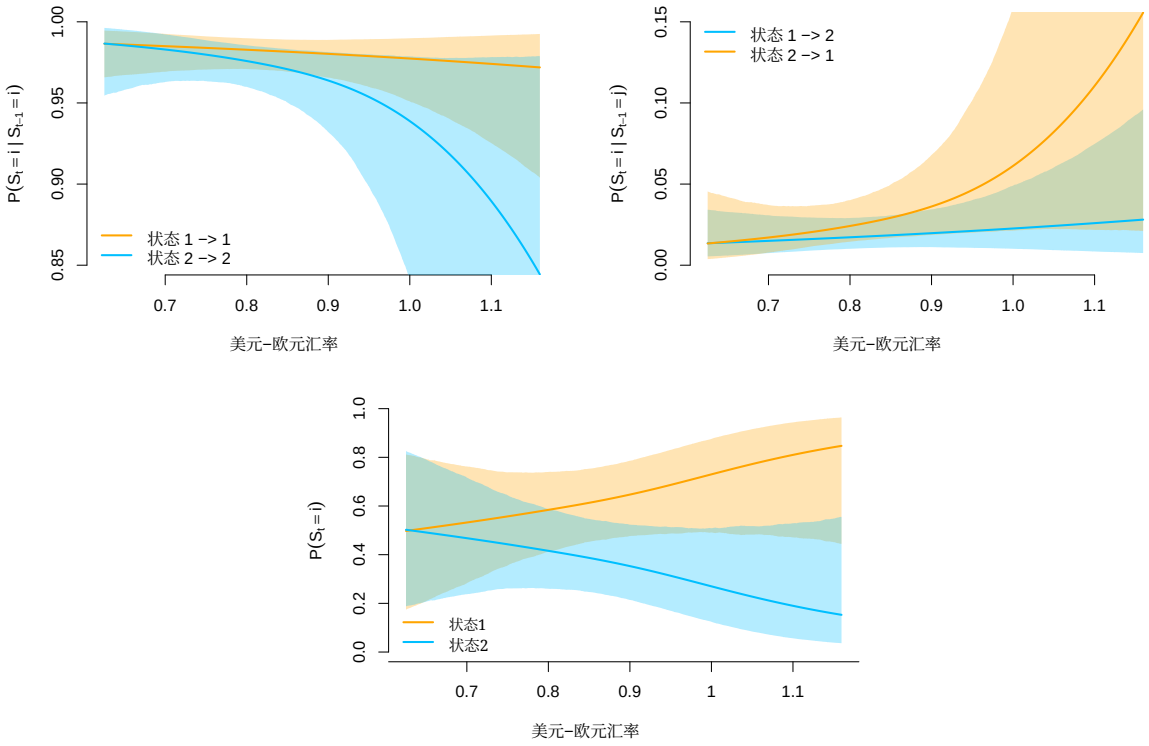

图9：估计的转移概率随汇率变化的情况。左上：保持在同一状态的概率。右上：不同状态之间切换的概率。下方：协变量相关的平稳状态概率随汇率变化的情况。阴影区域表示根据协变量相关转移概率的后验分布得到的95%点态置信区间。

图8分别按状态展示了石油价格对电价分布的条件均值（左上）、标准差（右上）和偏度（下方）的估计平滑效应。两种状态均表现出石油价格与电价之间正向且非线性的关联。整个协变量域内，状态1对应更高的条件均值，且估计平滑效应呈现出更明显的波动性。条件标准差在不同状态之间也存在明显差异。当 $40 \leq$ 石油价格$_t < 60$ 时，波动率升高。这一模式与解码状态序列一致，其中最高的条件方差与处于状态1的高概率相吻合。对偏度参数的影响取决于状态，并且呈现非线性。在低至中等石油价格下，两种状态的偏度均不断增加，其中状态1的偏度值始终更大，表明电价分布的右尾更加明显。随着石油价格进一步上升，状态1中的偏度效应趋于平稳并逐渐下降，而状态2则转向负偏度，这表明在较高石油价格水平下，分布转向更厚的下尾。

图9展示了估计的转移概率随美元—欧元汇率变化的情况。左图显示保持当前状态的概率，右图则显示在不同状态之间切换的概率。阴影区域表示根据估计的转移模型得出的95%置信区间。随着

汇率上升，两种状态的持续性均有所下降。然而，从状态2转移到状态1的概率上升，因此在较高汇率水平下进入状态1的似然更高。

协变量依赖的平稳分布（下方面板）显示，随着汇率上升，状态分配发生了明显变化。汇率上升时，状态1的平稳概率稳步增加，而处于状态2的概率则下降。显然，图9表明，汇率越高，状态1的持续性越强，这意味着在货币贬值期间，能源价格处于较高水平的时期会更加稳定。

这种模式与解码状态序列（见图7）相符，该序列将欧元水平相对较高的时期归为状态1。状态1的持续性增强，表明欧元升值时，较高的电价往往更加稳定，这可能是因为能源商品的进口成本降低。

为评估拟合模型的充分性，我们检查了图10所示的伪残差诊断结果。这些诊断的构建遵循第4.1节概述的相同原则。直方图和QQ图显示，伪残差与标准正态分布高度吻合，这表明偏态正态状态依赖分布能够对边际行为进行校准良好的描述。自相关函数显示伪残差仍存在序列依赖，这在未显式纳入自回归成分的情况下是预期的。

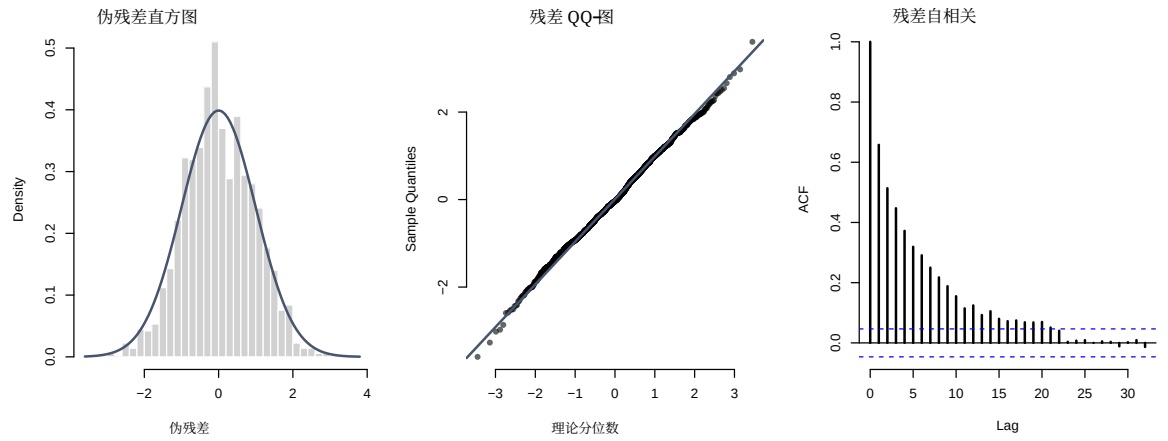

图10：拟合的能源价格模型的残差诊断。左：伪残差直方图，并叠加标准正态密度，显示经验分布与理论分布高度吻合。中：QQ图显示伪残差与标准正态分位数高度吻合，表明条件响应分布校准良好。右：伪残差的自相关函数显示仍存在残差自相关；这是预期之中的，因为模型未纳入用于捕捉时间依赖的显式自回归成分。

5 讨论

我们提出了一种非齐次 MS-GAMLSS 框架，用于对具有状态依赖分布参数的时间序列建模，其中隐状态过程和状态依赖过程都可能受协变量影响。将协变量纳入转移过程，有助于推断驱动状态转换动态的因素。具体而言，估计得到的转移概率表明，解释变量的变化如何影响潜在状态的持续性或状态间的转换。这进而有助于深入了解数据生成过程发生结构变化的内在机制。模拟实验证实，所提出的模型能够成功识别状态依赖参数和转移概率中的状态依赖协变量效应。实证应用进一步展示了这些效应的经济意义。

尽管实证示例依赖于正态和偏正态响应分布，但 GAMLSS 框架能够适应广泛的分布族。这些分布族包括连续数据、计数数据和二元数据的分布，以及在金融和经济应用中经常出现的偏态分布、厚尾分布和零膨胀分布。这种灵活性使所提出的框架不仅适用于本文考察的示例，也适用于多样的数据结构和典型特征。未来的研究可以利用这些分布扩展，使模型更贴合实证数据的具体属性。

除一般的分布建模之外，该框架还非常适合估计风险价值和预期短缺等风险度量指标。这些指标基于分位数而非期望，因此聚焦于分布尾部——这一点在金融应用中尤为重要。准确的尾部估计颇具挑战性，因为金融收益率序列通常呈现厚尾、偏度以及其他偏离正态性的特征。MS-GAMLSS 允许所有分布参数随状态和协变量变化，从而有望改善变化市场条件下的尾部估计。尽管这并非本研究的主要重点，但未来的工作可以探讨在需要完整分布预测的情境下，该模型是否适用于短期风险预测（Rockafellar 和 Uryasev，2002年；Acerbi 和 Tasche，2002年）。

进一步的扩展可以放宽有限个离散状态这一假设。在某些应用中，尤其是在金融时间序列领域，连续潜在状态过程（如状态空间模型中的过程）可能更为合适。允许不同状态之间连续转移，可以更平滑地适应潜在市场状况的渐进变化，也是未来研究的一个有前景的方向。

附录A 标准化前向变量

由于最大化过程中出现技术问题，例如数值下溢或上溢，面对较长的时间序列T时，似然可能变得极小或极大。因此，在 R等软件中，似然可能被舍入为零（下溢），或返回为 $+\infty$（上溢）。这是因为似然是通过矩阵乘积计算的，其中每个前向变量 $\boldsymbol{\alpha}_t$ 都会递归相乘，导致数值呈指数级增大或减小。为解决这一问题，我们采用缩放策略，引入标准化前向变量，并对其进行归一化以避免数值不稳定。令 $\boldsymbol{\phi}_t = (\phi_t(1)\ \phi_t(2) \ldots, \phi_t(N))$ 表示时间 t的标准化前向变量，定义如下

$$\boldsymbol{\phi}_t = \frac{\boldsymbol{\alpha}_t}{\boldsymbol{\alpha}_t \mathbf{1}^\top} = \frac{\boldsymbol{\alpha}_t}{\sum_{j=1}^{N} \alpha_t(j)}.$$

这种缩放确保标准化前向变量之和等于一，即$\sum_{j=1}^{N} \phi_t(j) = 1$，从而避免数值变得任意大或任意小。随后将 $\boldsymbol{\phi}_t$ 的递推更新为

$$\boldsymbol{\phi}_t = \frac{\boldsymbol{\phi}_{t-1} \boldsymbol{\Gamma} \mathbf{P}(y_t)}{\boldsymbol{\phi}_{t-1} \boldsymbol{\Gamma} \mathbf{P}(y_t) \mathbf{1}^\top},$$

其中$\mathbf{P}(y_t)$是时刻 t的状态依赖响应密度对角矩阵，而 $\boldsymbol{\Gamma}$是转移概率矩阵。

为了用标准化前向变量计算对数似然 $\ell(\boldsymbol{\theta})$，我们首先注意到，似然可以表示为各缩放值的乘积：

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{t=1}^{T} \frac{\boldsymbol{\alpha}_t \mathbf{1}^\top}{\boldsymbol{\alpha}_{t-1} \mathbf{1}^\top}.$$

取对数后，对数似然变为

$$\ell(\boldsymbol{\theta}) = \log\big(\mathcal{L}(\boldsymbol{\theta})\big) = \sum_{t=1}^{T} \log\left(\frac{\boldsymbol{\alpha}_t \mathbf{1}^\top}{\boldsymbol{\alpha}_{t-1} \mathbf{1}^\top}\right),$$

化简后得到

$$\ell(\boldsymbol{\theta}) = \log\big(\boldsymbol{\delta}^{(1)} \mathbf{P}(y_1) \mathbf{1}^\top\big) + \sum_{t=2}^{T} \log\big(\boldsymbol{\phi}_{t-1} \boldsymbol{\Gamma} \mathbf{P}(y_t) \mathbf{1}^\top\big).$$

参考文献

Carlo Acerbi 和 Dirk Tasche（2002年）。“预期短缺：风险价值的自然一致替代方案”。载于：*Economic Notes* 31.2，第379–388页。 DOI: 10.1111/1468-0300.00091。 URL: https://arxiv.org/abs/cond-mat/0105191。

Adam, Timo、Andreas Mayr 和 Thomas Kneib（2022）。“马尔可夫切换广义加性模型中位置、尺度和形状的梯度提升”。载于：*Econometricsand Statistics* 22，第3–16页。 DOI: 10.1016/j.ecosta.2022.05.003。

Balzer, Michael 等（2025）。“狄利克雷回归模型的梯度提升”。载于：AStA Advances in Statistical Analysis，第1–41页。 DOI: 10.1007/s10182-025-00526-5。英格兰银行（2017）。“如虎在侧：利用贷款级数据估计英国抵押贷款市场的脆弱性”。载于：工作人员工作论文第703号。 URL: https://www.bankofengland.co.uk/working-paper/2017/a-tiger-by-the-tail。Boor, Carl de（1978）。《样条实用指南》。纽约：施普林格。 ISBN: 978-1-4612- 6335-7。 DOI: 10.1007/978-1-4612-6333-3。德国汉莎航空股份公司（2024）。金融风险对冲。https://investor-relations.lufthansagroup.com/en/corporate-facts/financial-strategy-value-based-management/maintain-financial-stabilty/hedging-of-financial-risks.html。访问日期：2025-08-11。Eilers, Paul H. C. 和 Brian D. Marx（1996）。“使用B样条和惩罚项进行灵活平滑”。载于：Statistical Science 11.2，第89–121页。 DOI: 10.1214/ss/1038425655。Estrella, Arturo 和 Gikas A. Hardouvelis（1991）。“期限结构作为实际经济活动的预测指标”。载于：The Journal of Finance 46.2，第555–576页。 DOI: 10.1111/j.1540-6261.1991.tb02691.x。Foroni, Beatrice、Luca Merlo 和 Lea Petrella（2024）。“用于分析加密货币收益的期望分位数隐马尔可夫回归模型”。载于：Statistics and Computing 34.2，第66页。 DOI: 10.1007/s11222-023-10377-2。 URL: https://doi.org/10.1007/s11222-023-10377-2。Goldfeld, Stephen M. 和 Richard E. Quandt（1973）。“用于切换回归的马尔可夫模型”。载于：Journal of Econometrics 1.1，第3–15页。 DOI: 10.1016/0304-4076(73)90002-X。Hambuckers, Julien 等（2018）。“复合泊松过程的马尔可夫切换广义加性模型及其在操作损失模型中的应用”。载于：Quantitative Finance 18.10，第1679–1698页。 DOI: 10.1080/14697688.2017.1417625。 URL:https://doi.org/10.1080/14697688.2017.1417625。Hamilton, James D.（1989）。“非平稳时间序列与商业周期经济分析的新方法”。载于：Econometrica 57.2，第357–384页。 DOI: 10.2307/1912559。Hossain, Abu 等（2016）。“比例响应变量的百分位数估计”。载于：Statistics in Medicine 35.6，第895–904页。 DOI: 10.1002/sim.6748。Hu, Wenli、Brett A. Swanson 和 Gillian Z. Heller（2015）。“语音可懂度测试的统计分析方法”。载于：PLoS ONE 10.7。 DOI: 10.1371/journal.pone.0132409。国际货币基金组织（2015）。《美国金融部门评估规划压力测试——技术说明》。 URL: https://www.imf.org/external/pubs/ft/scr/2015/cr15173.pdf。Koslik, Jan-Ole（2024）。“通过准限制最大似然为非参数马尔可夫切换模型高效选择平滑度”。载于：arXiv预印本 arXiv:2411.11498。—（2025）。LaMa：隐马尔可夫模型的快速数值最大似然估计。R软件包版本2.0.5。 URL: https://CRAN.R-project.org/package=LaMa。Kristensen, Kasper（2025）。RTMB：“R”与TMB的绑定。R软件包版本1.7。 URL: https://CRAN.R-project.org/package=RTMB。Langrock, Roland、Timo Adam 等（2018）。“一般状态切换模型中基于样条的非参数推断”。载于：Statistica Neerlandica 72.3，第179–200页。 DOI:10.1111/stan.12151。

Langrock, Roland, Thomas Kneib 等（2017年）。《马尔可夫切换广义加性模型》。载于：《Statistics and Computing》27.1，第259–270页。 DOI: 10.1007/s11222-015-9620-3。Lystig, Theodore C. 和 James P. Hughes（2002年）。《隐马尔可夫模型观测信息矩阵的精确计算》。载于：《计算与图形统计学杂志》11.3，第678–689页。 DOI: 10.1198/106186002402。Regis, Renan O. 等（2023年）。《巴西金融市场中的资产定价：五因子 GAMLSS 建模》。载于：《实证经济学》64.5，第2373–2409页。 DOI: 10.1007/s00181-022-02316-3。Rigby, Robert A. 和 D. Mikis Stasinopoulos（2005年）。《位置、尺度和形状广义加性模型》。载于：《皇家统计学会杂志：C辑（应用统计）》54.3，第507–554页。 DOI: 10.1111/j.1467-9876.2005.00510.x。Rockafellar, R. Tyrrell 和 Stanislav P. Uryasev（2002年）。《一般损失分布的条件风险价值》。载于：《银行与金融杂志》26.7，第1443–1471页。DOI: 10.1016/S0378-4266(02)00271-6。Sanchez, J.、M.P. Mu˜noz 和 S. Fontdecaba（2009年）。《在 R 中估计马尔可夫切换回归模型：在西班牙能源价格建模中的应用》。载于：第173页。Sanchez-Espigares, Josep A. 和 Alberto Lopez-Moreno（2021年）。MSwM：拟合马尔可夫切换模型。R 软件包版本 1.5，CRAN，网址https://CRAN.R-project.org/package=MSwM。Stasinopoulos, Mikis D. 等（2024年）。《位置、尺度和形状广义加性模型：分布回归方法及其应用》。剑桥：剑桥大学出版社。 ISBN: 9781009410076。 DOI: 10.1017/9781009410076。 URL: https://doi.org/10.1017/9781009410076。Tong, Edward（2020年）。推荐信，美国银行。2020年7月14日。Ulrich, Matthias 等（2021年）。《电子杂货配送需求预测的分布回归》。载于：《欧洲运筹学杂志》294.3，第831–842页。 DOI: 10.1016/j.ejor.2021.01.022。Villarini, Gabriele 等（2009年）。《城市排水流域非平稳年最大值记录的洪水频率分析》。载于：《水资源进展》32.8，第1255–1266页。 ISSN: 0309-1708。 DOI: 10.1016/j.advwatres.2009.05.003。Viterbi, Andrew J.（1967年）。《卷积码的误差界及渐近最优译码算法》。载于：《IEEE信息论汇刊》13.2，第260–269页。 DOI: 10.1109/TIT.1967.1054010。Wattanawongwan, Suttisak 等（2023年）。《使用 GAMLSS 框架的信用卡违约风险暴露混合模型》。载于：《国际预测学杂志》39.1，第503–518页。 DOI: 10.1016/j.ijforecast.2021.12.014。Wood, Simon N.（2017年）。《广义加性模型：使用 R 入门》。第2版。Chapman and Hall/CRC。 ISBN: 9781498728331。 DOI: 10.1201/9781315370279。 URL: https://doi.org/10.1201/9781315370279。Zucchini, Walter、Iain L. MacDonald 和 Roland Langrock（2016年）。《时间序列隐马尔可夫模型：使用 R 入门》。第2版，第150卷。统计学与应用概率专著丛书。CRC出版社。 ISBN: 9781482253832。 DOI: 10.1201/9781315373744。 URL: https://doi.org/10.1201/9781315373744。