25

Kernal Functions in Gaussian Process

核函数

高斯过程中的核函数，先验协方差矩阵

人类的历史，本质上是思想的历史。

Human history is, in essence, a history of ideas.

—— 赫伯特·乔治·威尔斯 (Herbert George Wells) | 英国小说家和历史学家 | 1866 ~ 1946

- xxxxx
- xxxxxx
- xxxxxx
- xxxxxx

25.1 核函数

核函数 (kernel function) 是机器学习中一个常用概念。我们会在高斯过程 (Gaussian Process, GP)、支持向量机 (Support Vector Machine, SVM)、核主成分分析 (Kernel Principal Component Analysis, KPCA) 中用到核函数。本章则侧重高斯过程中的核函数。《机器学习》将专门介绍支持向量机中的核技巧 (kernel trick)，以及核主成分分析中的核函数。

经过上一章的学习，大家已经清楚高斯过程是一种用于回归和分类的非参数模型，它通过对输入空间中数据点之间的相似性进行建模，从而实现对输出的推断。

核函数在高斯过程中的作用是定义输入空间中数据点之间的相似性或相关性。以上一章介绍过的高斯核为例，核函数决定了在输出空间中，两个数据点对应的预测值之间的协方差。

如图 1 所示，如果两个数据点 (x_i, x_j) 在输入空间中相似，它们对应的输出值就有更高的协方差 (深蓝色)，反之则有较低的协方差 (浅蓝色)。注意，上述描述适用于高斯核协方差矩阵，并不能描述其他常见核函数。表 1 总结高斯过程中常见的核函数。

下面，我们展开聊聊线性核、高斯核、周期核。本章最后还会介绍核函数叠加。

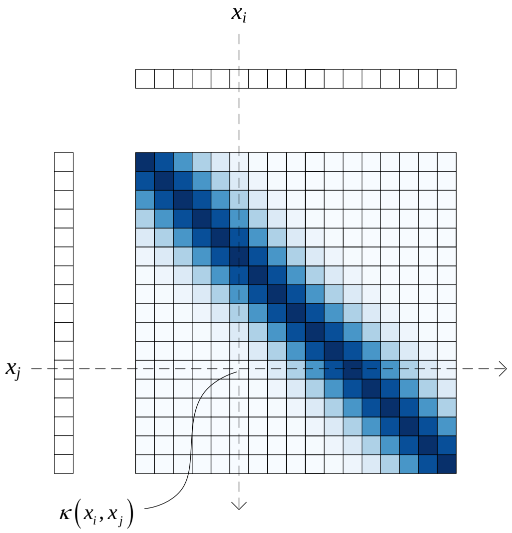
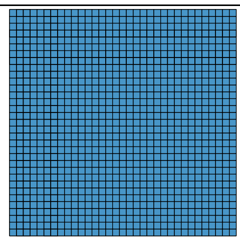
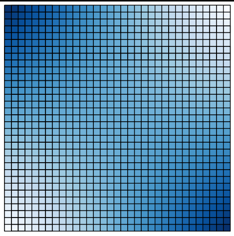
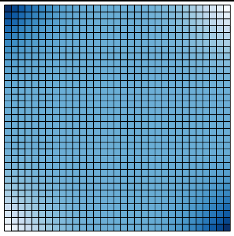
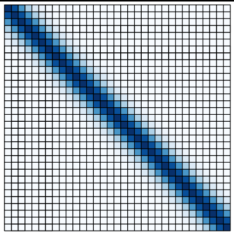
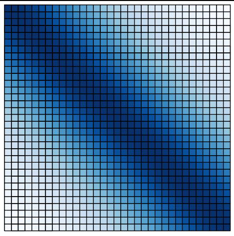
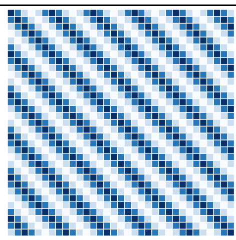


图 1. 高斯过程的高斯核协方差矩阵

表 1. 高斯过程中常用的核函数

核函数	常见形式	先验协方差矩阵
常数核 (constant kernel)	$\kappa(x_i, x_j) = c$	

线性核 (linear kernel)	$\kappa\left(x_i,x_j\right)=\sigma^2\left(x_i-c\right)\left(x_j-c\right)$ $\kappa\left(x_i,x_j\right)=\sigma_b^2+\sigma^2\left(x_i-c\right)\left(x_j-c\right)$	
多项式核 (polynomial kernel)	$\kappa\left(x_i,x_j\right)=\left(\left(x_i-c\right)\left(x_j-c\right)+\text{offset}\right)^d$ $\kappa\left(x_i,x_j\right)=\left(\sigma^2\left(x_i-c\right)\left(x_j-c\right)+\text{offset}\right)^d$	
高斯核 (Gaussian kernel)	$\kappa\left(x_i,x_j\right)=\exp\left(-\frac{\left(x_i-x_j\right)^2}{2l^2}\right)$ $\kappa\left(x_i,x_j\right)=\sigma^2\exp\left(-\frac{\left(x_i-x_j\right)^2}{2l^2}\right)$	
有理二次核 (rational quadratic kernel)	$\kappa\left(x_i,x_j\right)=\exp\left(1+\frac{\left(x_i-x_j\right)^2}{2\alpha l^2}\right)^{-\alpha}$ $\kappa\left(x_i,x_j\right)=\sigma^2\exp\left(1+\frac{\left(x_i-x_j\right)^2}{2\alpha l^2}\right)^{-\alpha}$	
周期核 (periodic kernel)	$\kappa\left(x_i,x_j\right)=\sigma^2\exp\left(-\frac{2\sin^2\left(\frac{\pi}{p}\left\|x_i-x_j\right\|\right)}{l^2}\right)$ $\kappa\left(x_i,x_j\right)=\sigma^2\exp\left(-\frac{\sin^2\left(\frac{\pi}{p}\left\|x_i-x_j\right\|\right)}{2l^2}\right)$	

25.2 线性核

线性核是高斯过程中的一种核函数，也称为线性相似性函数。它是一种用于衡量输入数据点之间线性关系的核函数。本节采用的线性核函数的表达式为

$$\kappa\left(x_i,x_j\right)=\operatorname{cov}\left(y_i,y_j\right)=\sigma^2\left(x_i-c\right)\left(x_j-c\right) \tag{1}$$

其中，x_i 和 x_j 分别是输入空间中的两个数据点。图 2 所示为线性核先验协方差矩阵的曲面。

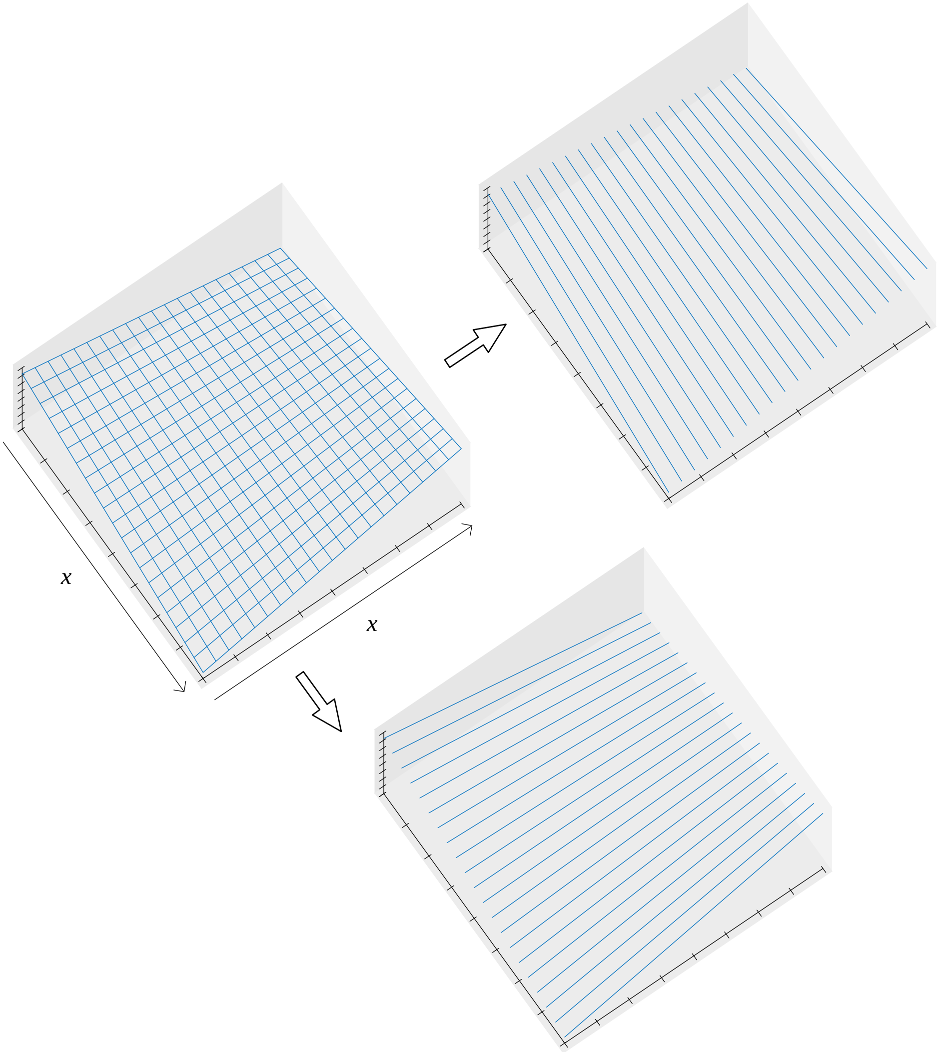


图 2. 无限大的先验协方差矩阵，线性核

图 3 所示为当 $c = 0$ 时，线性核参数 σ 对先验协方差和采样影响。注意，图 3 中不同热图子图的颜色映射取值范围不同。

线性核函数假设输入空间中的数据点之间存在线性关系，即输出值之间的相似性与输入数据点的线性组合有关。如果两个数据点在输入空间中更接近形成线性关系，它们对应的输出值在高斯过程中的协方差较高。

线性核函数在某些问题中很有效，特别是当数据呈现线性关系时。然而，对于非线性关系的数据，其他核函数如高斯核函数可能更适用，因为它们能够处理更复杂的数据结构。

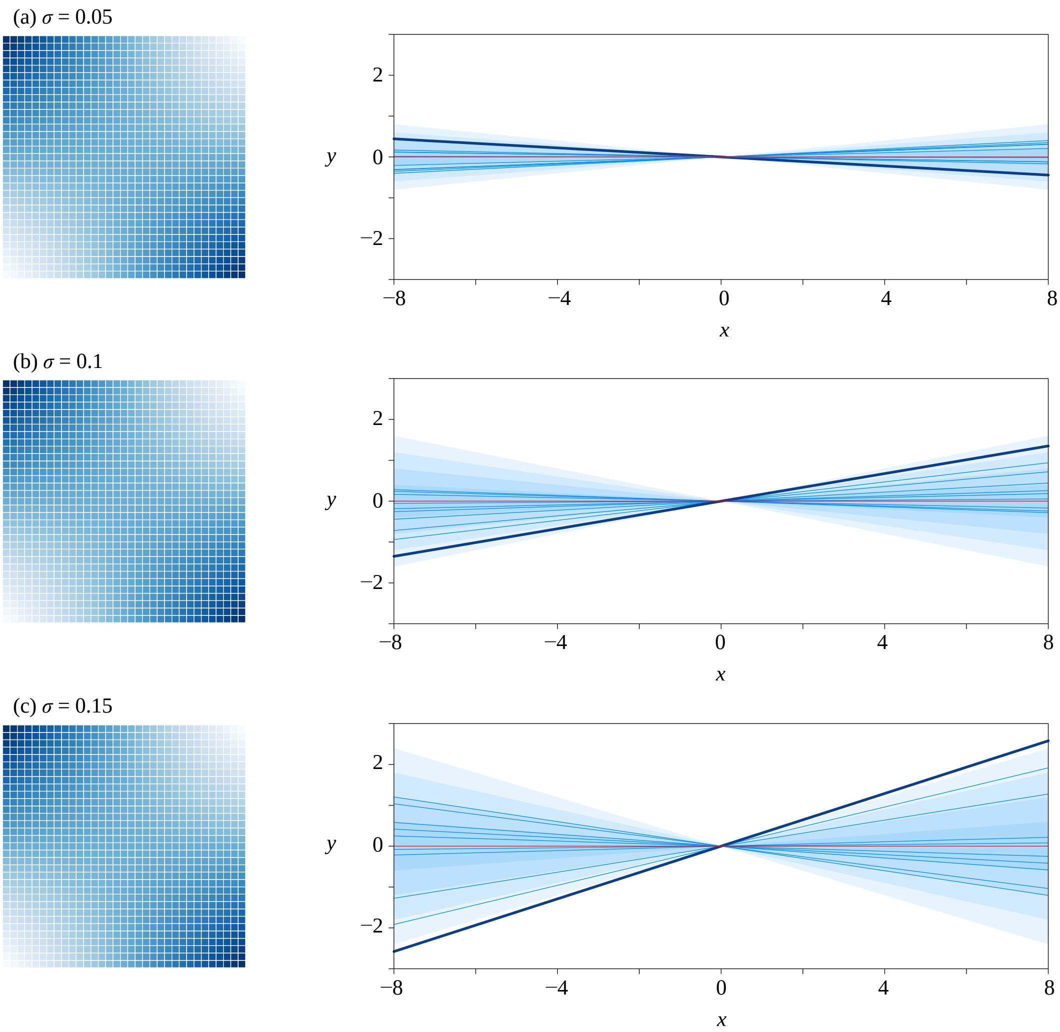


图 3. 线性核参数 σ 对先验协方差和采样影响

25.3 高斯核

高斯核是在高斯过程中常用的核函数之一，也称为径向基函数 (Radial Basis Function, RBF) 或指数二次核 (exponentiated quadratic kernel, squared exponential)。本节采用的高斯核的形式为

$$\kappa\left(x_i, x_j\right) = \operatorname{cov}\left(y_i, y_j\right) = \sigma^2 \exp\left(-\frac{\left(x_i - x_j\right)^2}{2l^2}\right) \tag{2}$$

其中，σ^2 为协方差矩阵的方差；l 是高斯核长度尺度度量参数。请大家翻阅上一章查看高斯核的先验协方差曲面。

前文提过，高斯核函数的作用是衡量两个输入点之间的相似性，当两点距离较近时，核函数的值较大，表示它们在函数空间中具有相似的输出；反之，距离较远时核函数的值较小，表示它们在输出上差异较大。

图 4 所示为当 $\sigma = 1$ 时高斯核参数 l 对先验协方差和采样影响。比较几幅子图，我们可以发现 l 越大，协方差矩阵中协方差相对更大 (临近点协同运动越强)，对应曲线越平滑。

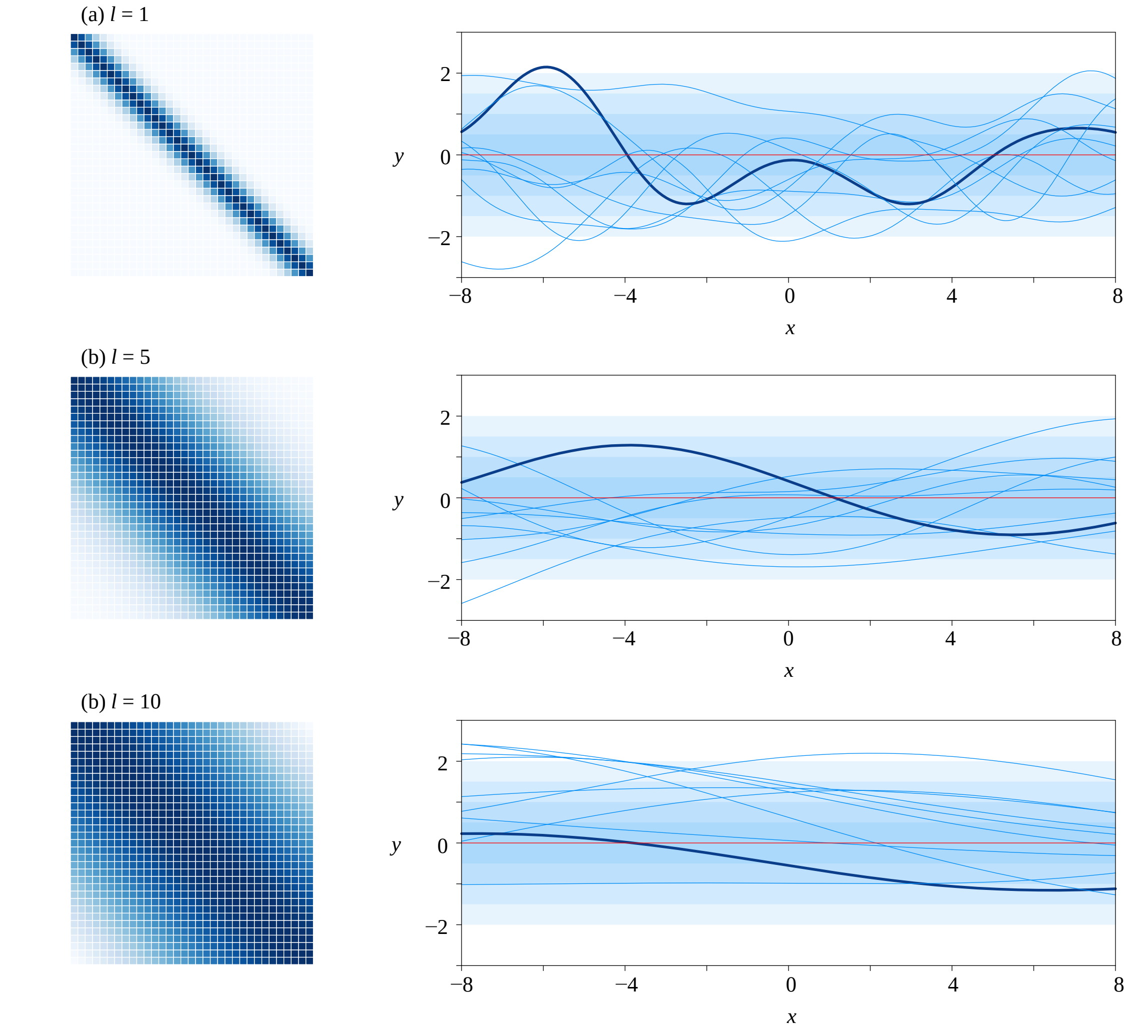


图 4. 高斯核参数 l 对先验协方差和采样影响，$\sigma = 1$

对于 (2)，我们发现 σ 也是高斯核参数之一。图 5 所示为 $l = 1$ 时高斯核参数 σ 对先验协方差和采样影响。

高斯核具有平滑性和无限可微性的特点，使其在建模各种复杂函数时表现出色。在高斯过程中，选择合适的核函数和参数是关键，它直接影响了模型对数据的拟合程度和泛化能力。

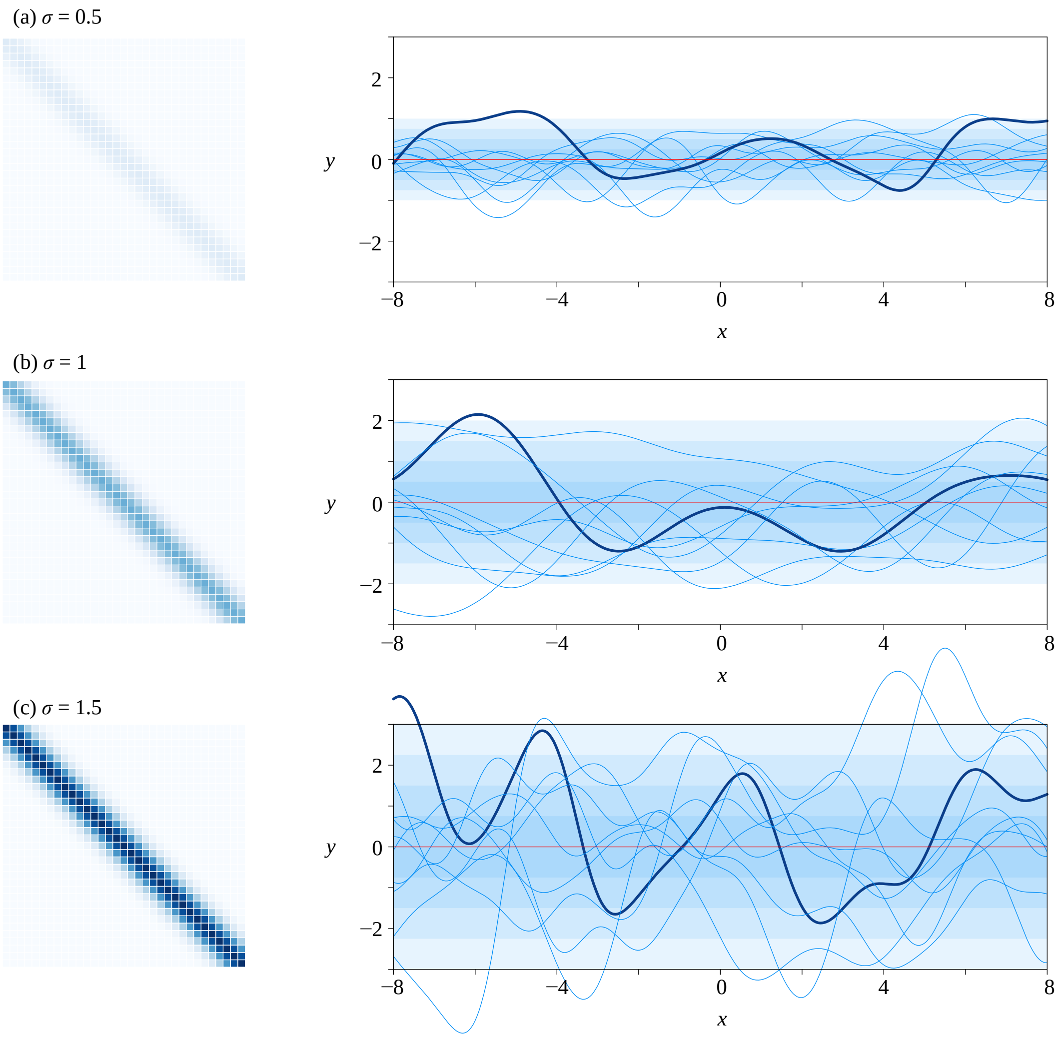


图 5. 高斯核参数 σ 对先验协方差和采样影响，$l = 1$

25.4 周期核

周期核是高斯过程中常用的核函数之一，它适用于描述具有周期性变化的数据。本节采用的周期核的形式为

$$\kappa\left(x_i, x_j\right) = \operatorname{cov}\left(y_i, y_j\right) = \sigma^2 \exp\left(-\frac{\sin^2\left(\frac{\pi}{p}\left|x_i - x_j\right|\right)}{2l^2}\right) \tag{3}$$

其中，p 影响周期核的周期，l 是高斯核的长度尺度参数。参数 l 对周期核的影响类似高斯核。

图 6 所示为周期核先验协方差曲面。

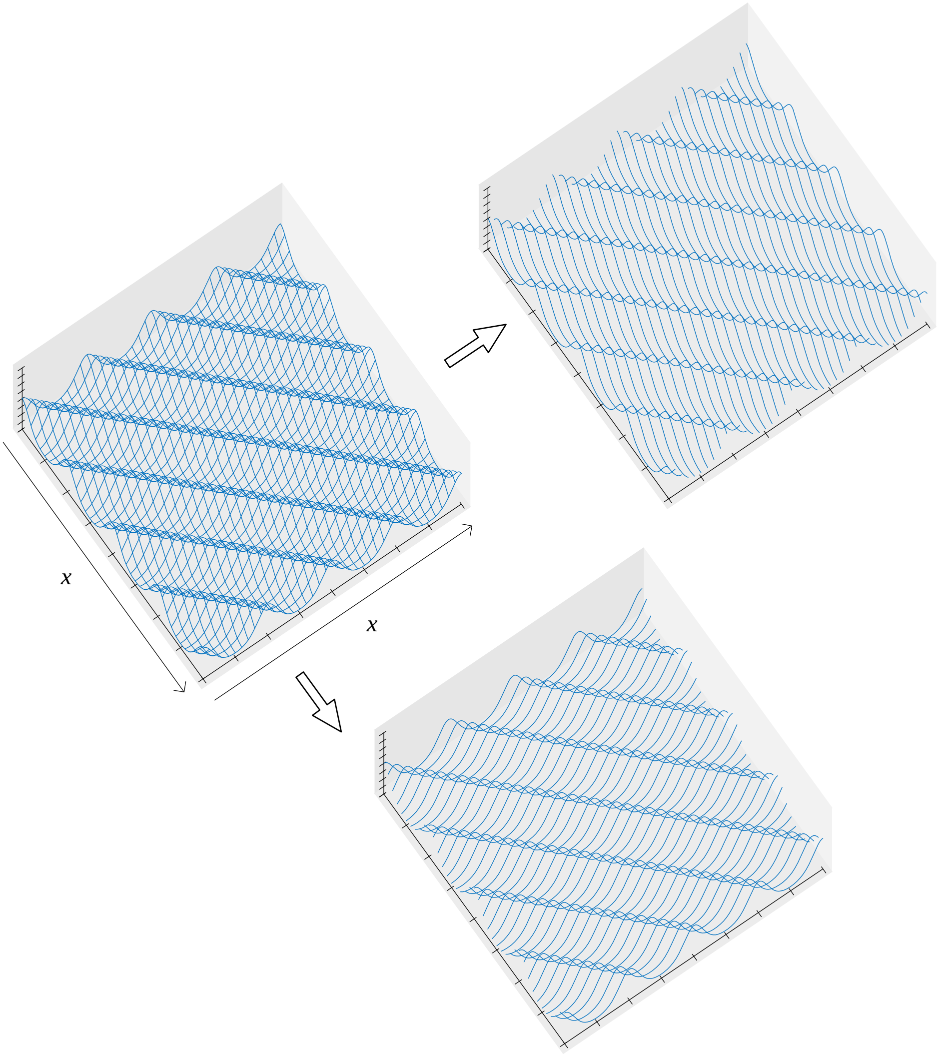


图 6. 无限大的先验协方差矩阵，周期核

下面着重介绍一下参数 p 周期核影响。

如图 7 所示，随着参数 p 增大，采样曲线的波动周期不断变长。

周期核的关键特点在于它引入了正弦函数，使得核函数对周期性变化非常敏感。当输入点在周期上相隔较短的距离时，核函数的值较大，表示这些点在函数空间中具有相似的输出。而当输入点在周期上相隔较远时，核函数的值较小，表示它们在输出上有较大的差异。

周期核常用于建模具有明显周期性结构的时间序列数据或周期性变化的信号。选择合适的周期和长度尺度参数是使用周期核的关键，这样可以使高斯过程模型更好地捕捉数据中的周期性模式。

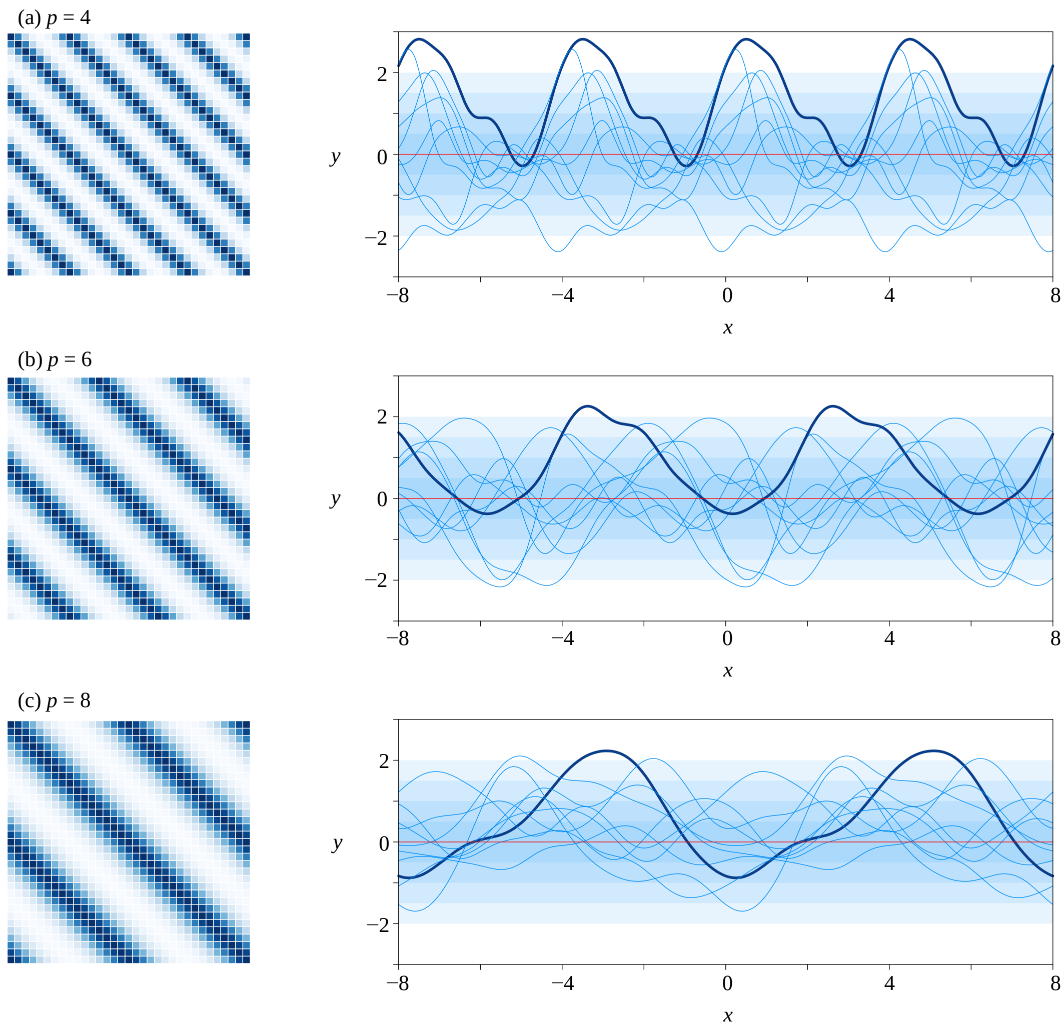


图 7. 高斯核参数 σ 对先验协方差和采样影响，$l = 1$

25.5 核函数的组合

在高斯过程中，我们还可以通过加法或乘法叠加不同核函数以便构建更复杂、更灵活的核函数。这种方式使得高斯过程模型能够更好地适应不同类型的数据模式。

比如，通过乘法获得两个核函数的乘积：

$$\kappa\left(x_i, x_j\right) = \operatorname{cov}\left(y_i, y_j\right) = \kappa_1\left(x_i, x_j\right) \cdot \kappa_2\left(x_i, x_j\right) \tag{4}$$

这种方式常用于组合两个核函数的优点，例如结合具有周期性和长度尺度的核函数，以适应同时存在周期性和趋势性的数据。当然，我们也可以将更多不同类型核函数通过乘积方式组合起来。

再比如，通过加法获得两个核函数的和：

$$\kappa\left(x_i, x_j\right) = \operatorname{cov}\left(y_i, y_j\right) = \kappa_1\left(x_i, x_j\right) + \kappa_2\left(x_i, x_j\right) \tag{5}$$

通过加法叠加，可以将两个核函数的特性相加，得到新的核函数。这种方式常用于处理数据中不同尺度的变化，例如同时存在高频和低频成分的数据。

下面举三个例子，用乘法组合两个不同的核函数。

图 8 所示为高斯核和线性核的乘积，即高斯核 × 线性核。图 9 所示为高斯核 × 线性核先验协方差矩阵及采样曲线。

线性核部分可以处理数据的线性趋势，而高斯核部分则引入了非线性特性，能够捕捉数据中的复杂模式和局部关系。在实际应用中，这种组合核函数常用于处理同时存在线性和非线性结构的数据。例如，当数据在整体上呈线性趋势，但在局部存在一些非线性的波动或变化时，使用线性核和高斯核的乘积可以更好地拟合数据。

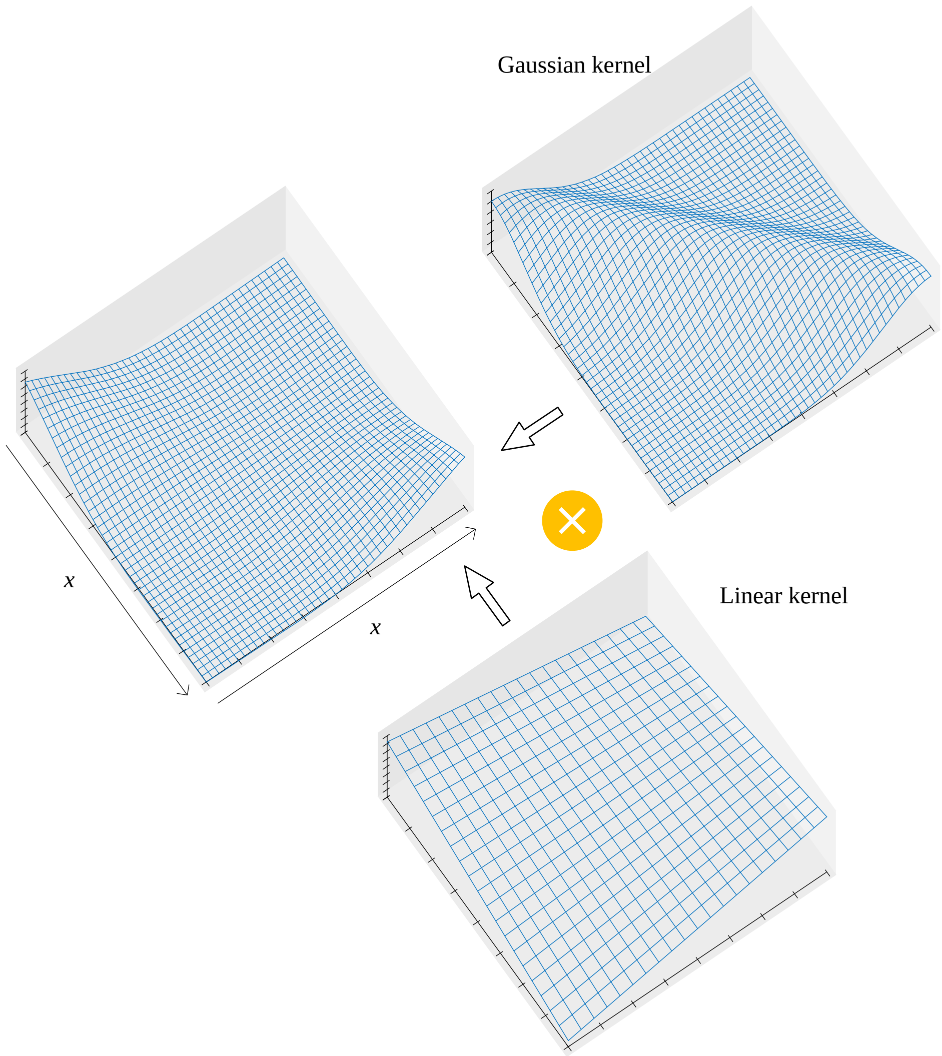


图 8. 高斯核和线性核的乘积

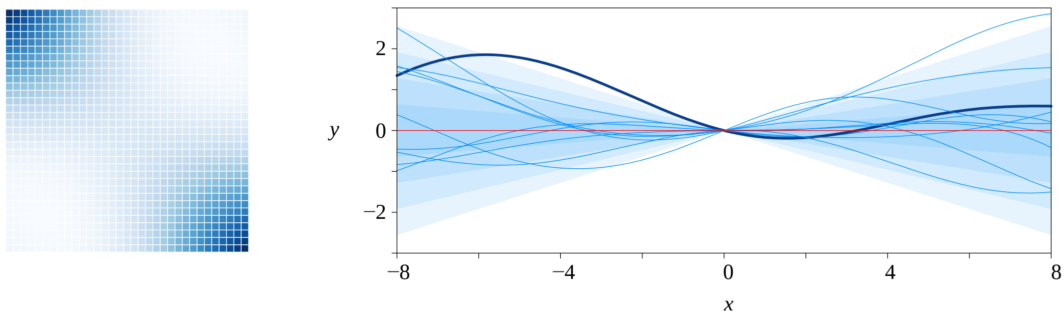


图 9. 高斯核和线性核的乘积，先验协方差矩阵和采样

图 10 所示为高斯核和周期核的乘积，即高斯核 × 周期核。图 11 所示为高斯核 × 周期核先验协方差矩阵及采样曲线。这样的组合核函数结合了周期性和非周期性的特性。周期核部分能够捕捉数据中的周期性结构，而高斯核部分引入了非周期性的平滑性，使模型对整体趋势有更好的拟合能力。

在实际应用中，这种组合核函数常用于建模具有明显周期性变化，同时又包含一些噪声或非周期性成分的数据。例如，对于时间序列数据，可能存在明显的季节性变化（周期性），同时受到其他因素的影响（非周期性）。通过将周期核和高斯核进行乘积，模型能够更全面地考虑这两种特性，提高对复杂数据模式的拟合能力。

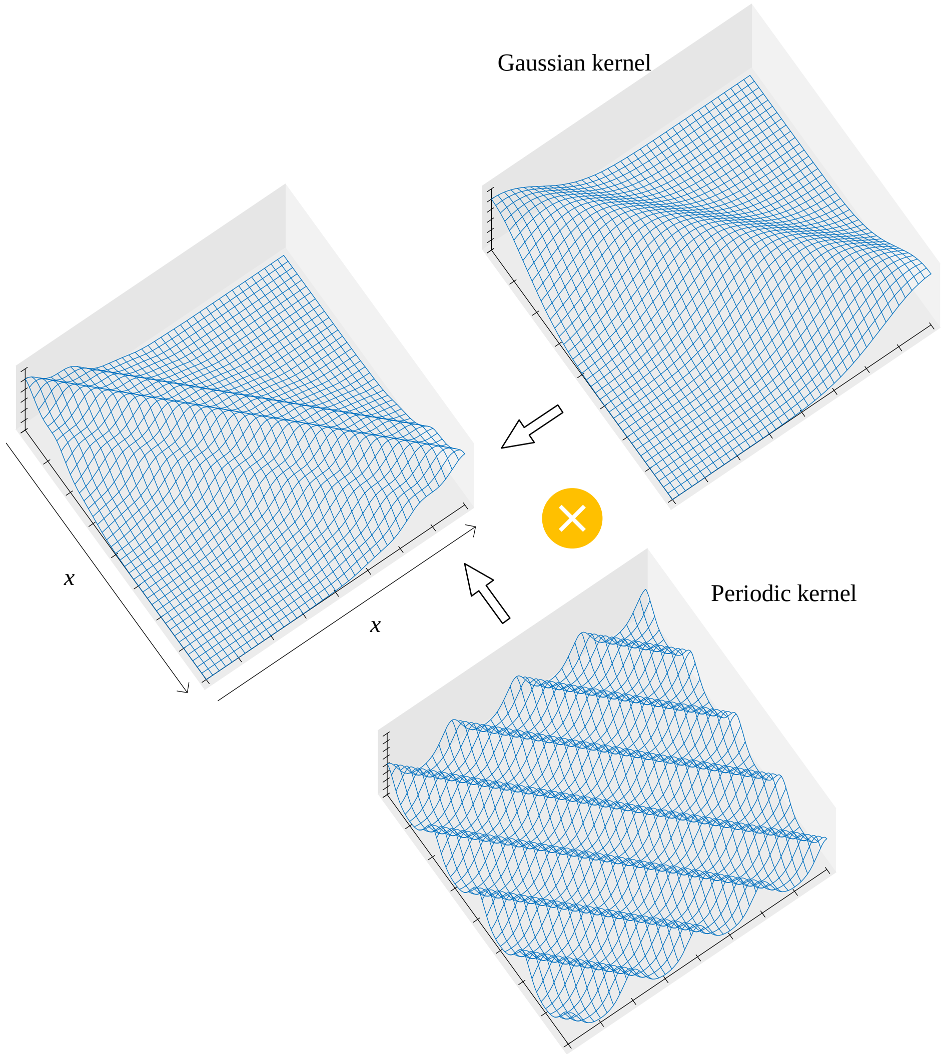


图 10. 高斯核和周期核的乘积

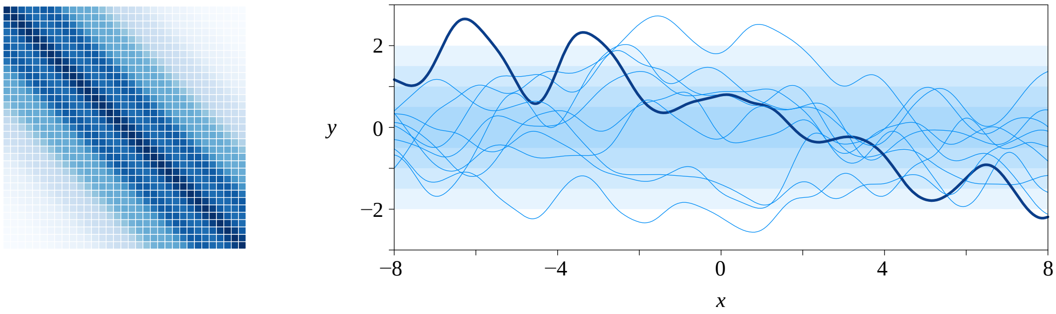


图 11. 高斯核和周期核的乘积，先验协方差矩阵和采样

图 12 所示为线性核和周期核的乘积，即线性核 × 周期核。图 13 所示为线性核 × 周期核先验协方差矩阵及采样曲线。这样的组合核函数同时包含了周期性和线性趋势的特性。周期核部分捕捉了数据中的周期性结构，而线性核部分用于处理数据的线性趋势。通过这种组合，模型可以更灵活地适应同时存在周期性和线性结构的数据。

在实际应用中，这种组合核函数常用于处理同时具有周期性和整体线性趋势的数据，例如时间序列数据中同时存在季节性变化和总体趋势的情况。

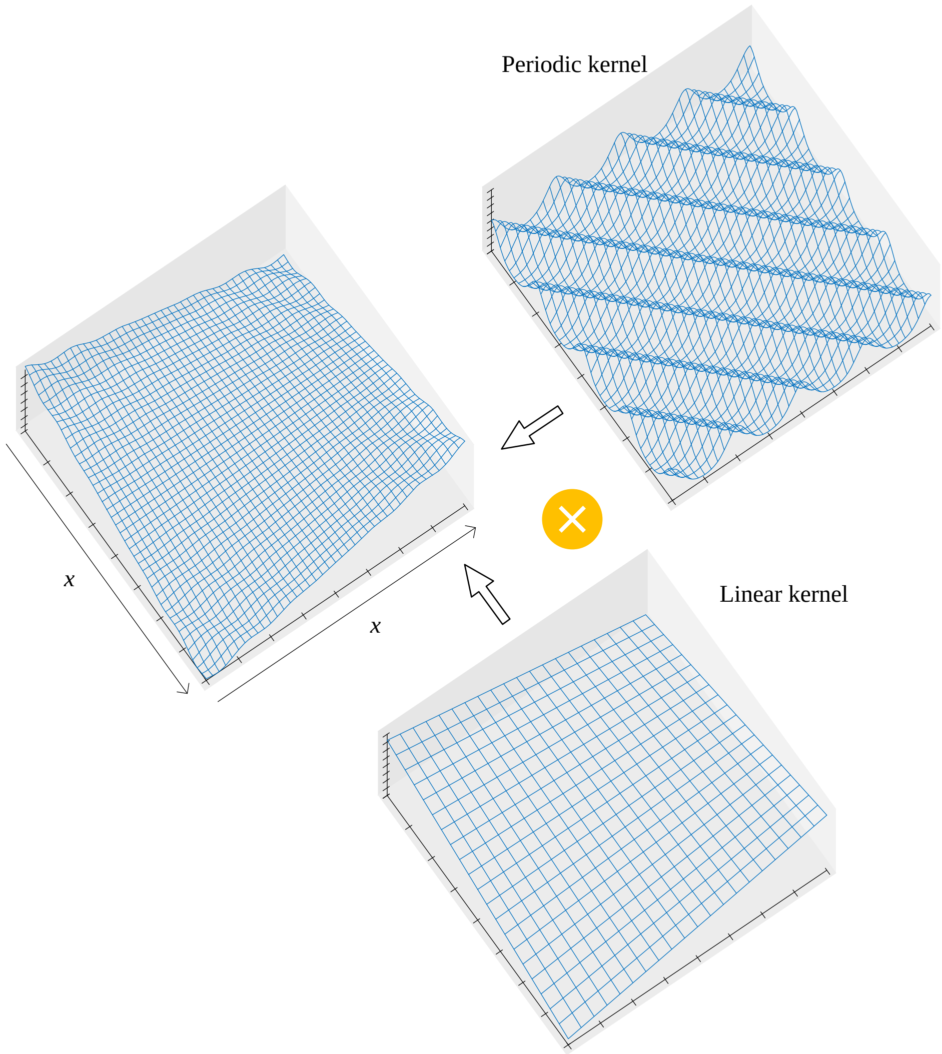


图 12. 线性核和周期核的乘积

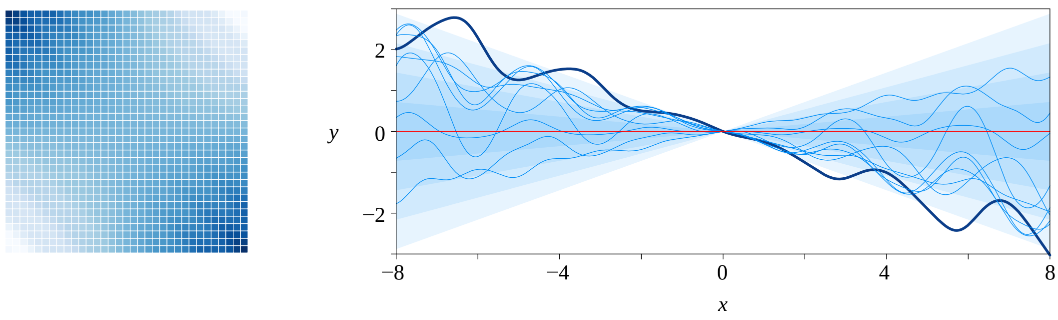


图 13. 线性核和周期核的乘积，先验协方差矩阵和采样

《机器学习》在讲解支持向量机核技巧时还会提到不同的核函数。此外，《机器学习》还会介绍如何用 Scikit-learn 库中高斯过程工具完成回归和分类。

大家想要深入学习高斯过程，请参考如下开源图书 *Gaussian Processes for Machine Learning*：

https://gaussianprocess.org/gpml/

这篇博士论文中专门介绍了不同核函数的叠加：

https://www.cs.toronto.edu/~duvenaud/thesis.pdf

作者认为下面这篇文章解释高斯过程做的交互设计最佳，这篇文章给了作者很多可视化方面的启发：

https://distill.pub/2019/visual-exploration-gaussian-processes/