سمینار درس

یادگیری ماشین

زمستان ۱۴۰۰

موضوع: مقدمه‌ای بر روش‌های طبقه‌بندی و رگرسیون مبتنی بر فرآیند گوسی

ارائه دهنده: .....

فهرست مطالب

- فرآیند گوسی
- رگرسیون فرآیند گوسی
- مثال رگرسیون گوسی
- طبقه‌بندی فرآیند گوسی
- مثال طبقه‌بندی گوسی
- هسته در فرآیند گوسی
- منابع

فرآیند گوسی(GP)

- فرآیند گوسی یک روش یادگیری با نظارت است.

❖ مزایا

۱. استفاده از درون یابی برای پیش‌بینی‌ها
۲. تصمیم‌گیری بر اساس فواصل اطمینان
۳. هسته‌های مشترک وخصوصی

❖ معایب

۱. استفاده از کل اطلاعات نمونه‌ها
۲. عملکرد ضعیف در فضاهای با ابعاد بالا

رگرسیون فرآیند گوسی(Gaussian Process Regression)

۱. محاسبه احتمال پیشین.

۲. محاسبه کوواریانس احتمال پیشین با استفاده از هسته.

۳. بهینه سازی هایپرپارامترهای هسته در طول برازش رگرسیون فرآیند گاوسی به‌وسیله حداکثرکردن احتمال شباهت حاشیه‌ای(LML).

✓ بهینه‌سازی مجدد به دلیل وجود چندین نقطه محلی در LML.

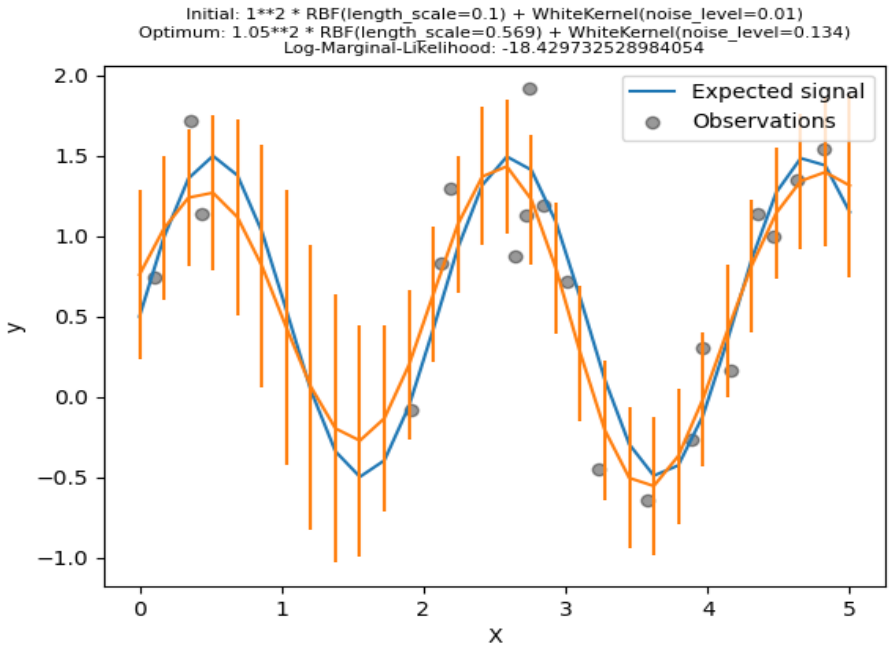
✓ تعیین سطح نویز به وسیله هسته سفید (White Kernel).

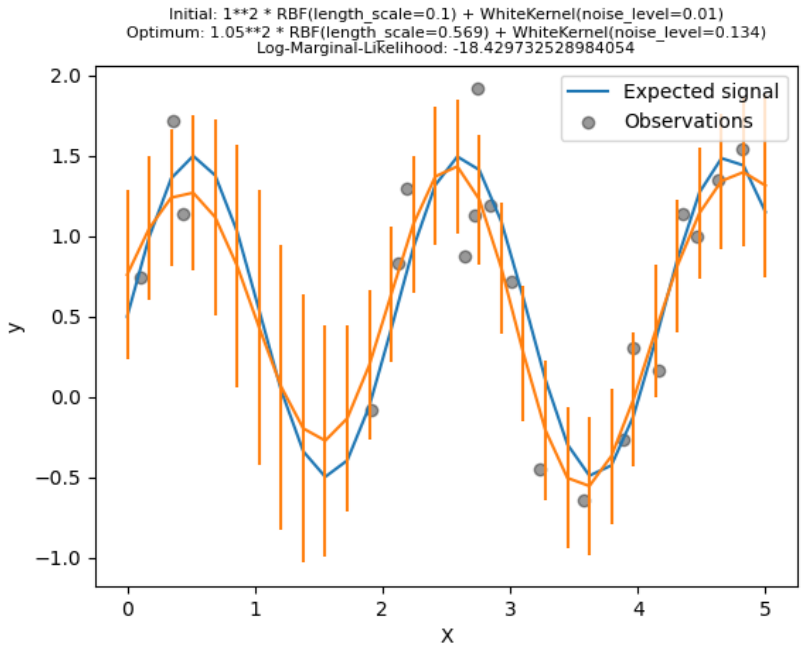
✓ مزیت استفاده از هسته سفید: امکان پیش‌بینی، بدون برازش احتمال پیشین.

مثال رگرسیون فرآیند گوسی

- ✓ رگرسیون فرآیند گاوسی با تخمین سطح نویز
- ✓ توانایی هسته سفید برای تخمین سطح نویز

بهینه‌سازی چندین بار برای مقداردهی‌های اولیه مختلف تکرار می‌شود.

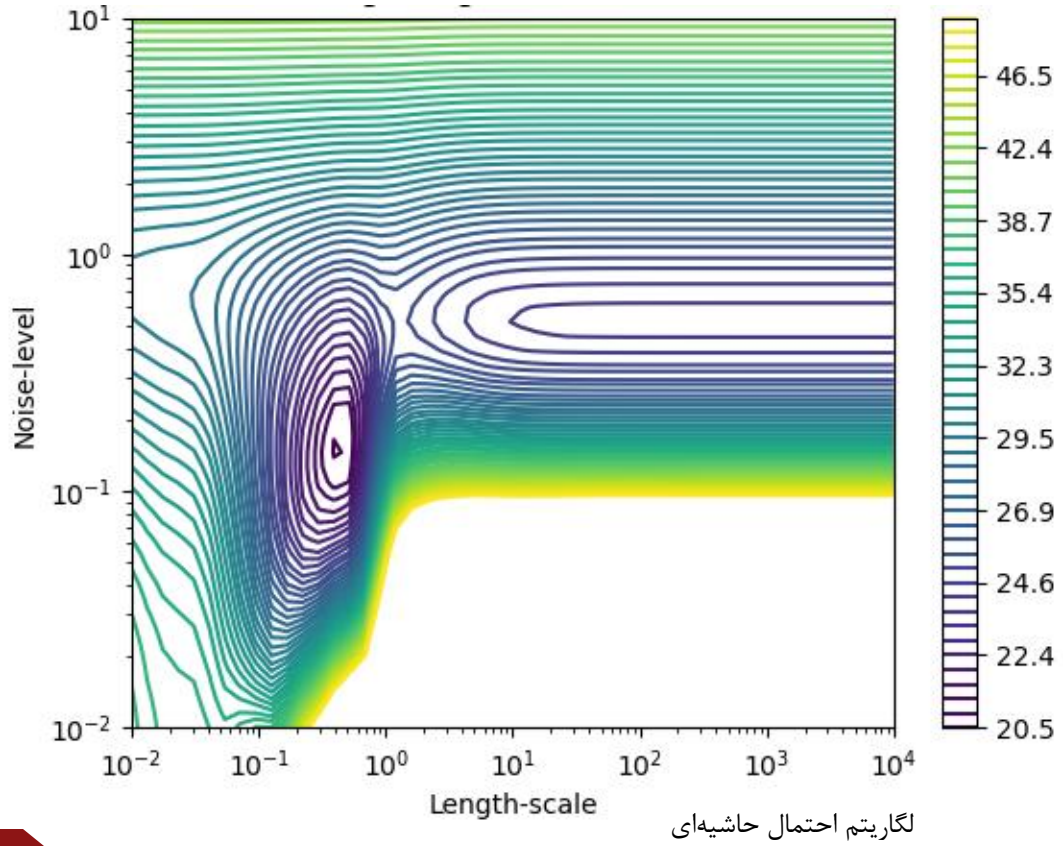




.(۲۰۲۲/۰۱)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html

مثال رگرسیون فرآیند گوسی

- رگرسیون فرآیند گاوسی با تخمین سطح نویز

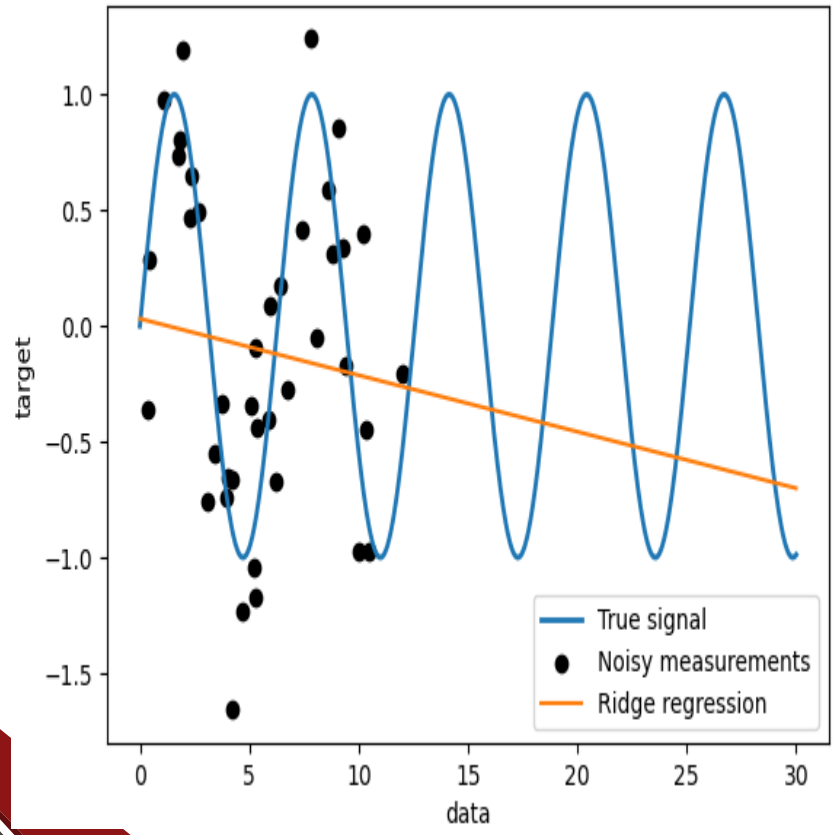


لگاریتم احتمال حاشیه‌ای

.(۲۰۲۲/۰۱)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html

مثال رگرسیون فرآیند گوسی

- سینوسی با دوره تناوب 2π
- اندازه‌گیری همراه با نویز است.
- فقط نمونه‌هایی از ابتدای سیگنال موجود است.



- **محدودیت یک مدل خطی مانند ریج هسته**

چنین رگرسیون‌ای داده‌ها را از بین می‌برد زیرا به‌اندازه کافی معنادار نیست.

.(۲۰۲۲) .'Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_compare_gpr_krr.html

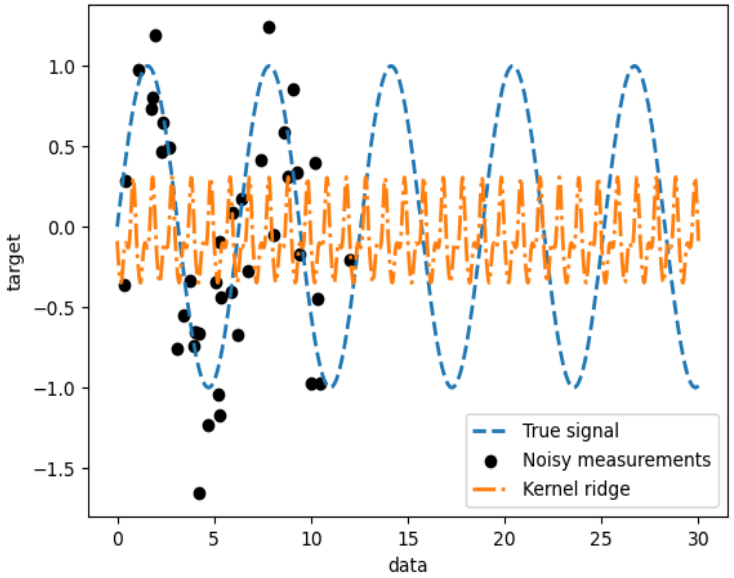
مثال رگرسیون فرآیند گوسی(ادامه)

گویا تر کردن مدل قبلی با تغییرات هسته امکان دارد.

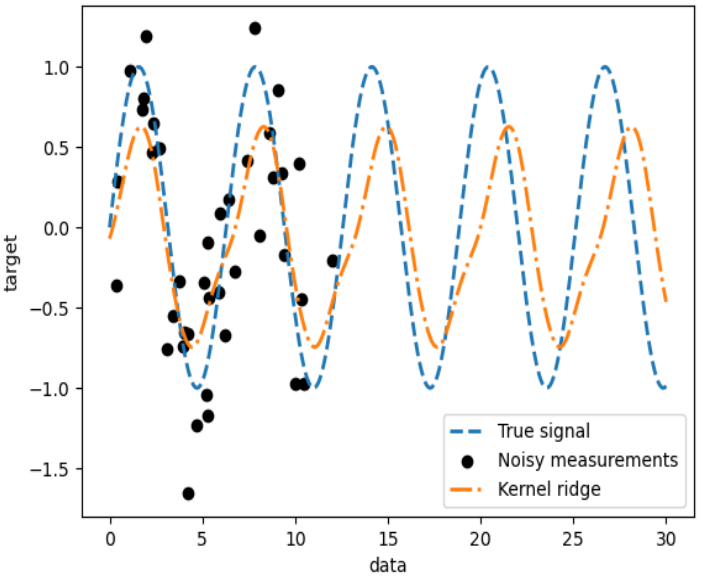
می‌توان هسته مربع سینوسی نمایی(Exp Sine Squared) استفاده کرد.

از یک جستجوی تصادفی برای تنظیم پارامترهای مختلف مدل ریج هسته استفاده می‌شود.

اما هنوز برخی از خطاها عمدتاً به دلیل نویز اضافه‌شده به مجموعه داده مشاهده می‌شود.



رگرسیون ریج هسته با هسته مربع سینوسی نمایی با استفاده از هایپرپارامترهای پیش‌فرض



رگرسیون ریج هسته با یک هسته مربع سینوسی نمایی با استفاده از هایپرپارامترهای تنظیم‌شده

.(۲۰۲۲) .'Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_compare_gpr_krr.html

مثال رگرسیون فرآیند گوسی(ادامه)

✓ **رگرسیون گوسی(GPR)**

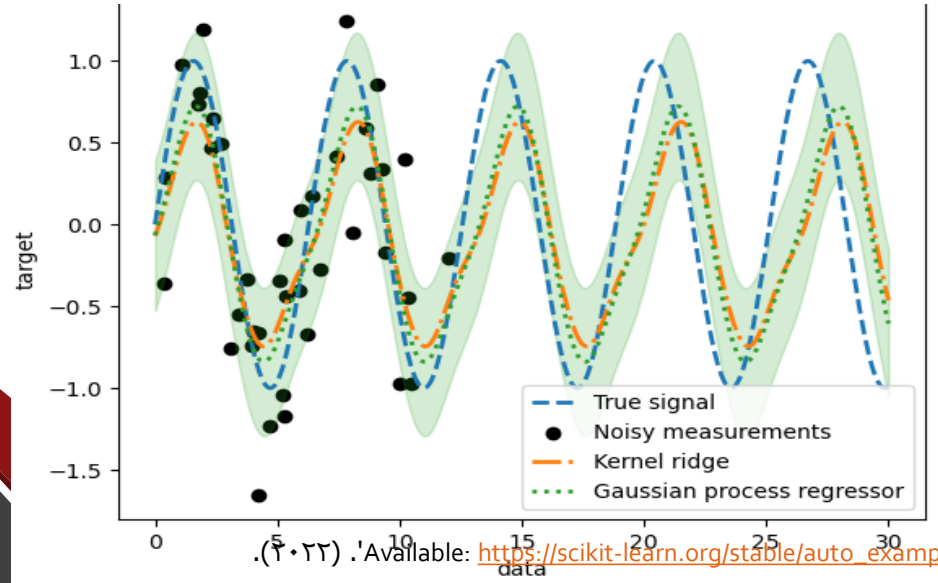
افزودن هسته سفید و محاسبه نویز داده‌ها.

✓ **مقایسه رگرسیون گوسی(GPR) و رگرسیون ریج هسته(Kernel Ridge Regression)**

هر دو روش مدل منطبقی بر تابع هدف دارند.

در GPRدوره تناوب برابر $6.28 * 2\pi$(نزدیک به دوره تناوب سیگنال پریودیک)و در KRR دوره تناوب $4 * \pi$.

به‌طور کلی برازشKRR سریعتر تر است اما مشکل Curse of dimensionality دارد.



مقایسه رگرسیون گوسی و رگرسیون ریج هسته

'Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_compare_gpr_krr.html. (۲۰۲۲).

مثال رگرسیون فرآیند گوسی(ادامه)

مقایسه رگرسیون گوسی(GPR) و رگرسیون ریج هسته(Kernel Ridge Regression)

شباهت KRR و GPR: یادگیری تابع هدف با استفاده از قانون هسته(Kernel trick).

تفاوت

KRR:

یادگیری یک تابع خطی بر اساس خطای میانگین مربع(Mean-squared error) با استفاده از هسته

GPR:

استفاده از هسته برای تعریف کوواریانس توزیع احتمال پیشین بر روی توابع هدف و استفاده از داده‌های آموزشی مشاهده‌شده برای تعریف یک تابع شباهت(Likelihood)

بر اساس قضیه بیز، یک توزیع پسین(Posterior distribution)گاوسی بر روی توابع هدف تعریف شده است که میانگین آن برای پیش‌بینی استفاده می‌شود.

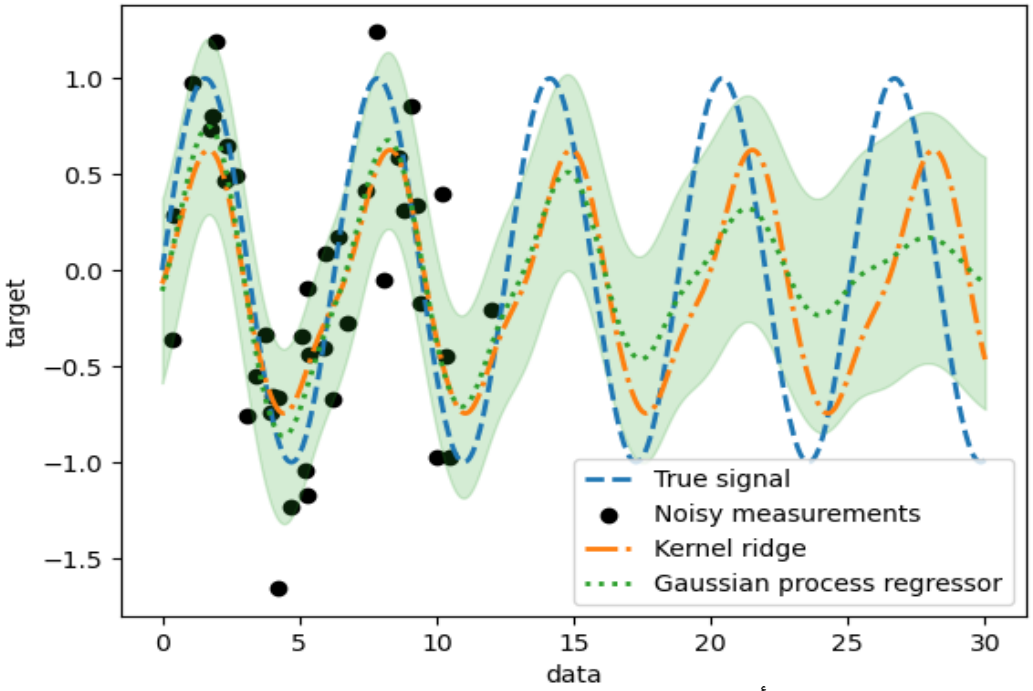
GPR: انتخاب هایپرپارامترهای هسته بر اساس افزایش گرادیان بر روی تابع احتمال حاشیه‌ای.

KRR: جستجوی شبکه‌ای(Grid search)را بر روی یک تابع هزینه(خطای میانگین مربع).

مثال رگرسیون فرآیند گوسی(ادامه)

فرآیند گاوسی اجازه می‌دهد تا هسته‌ها را باهم ترکیب کنیم.

هسته مربعی سینوسی نمایی + هسته تابعی بر مبنای شعاعی(Radial basis function kernel)



تأثیر استفاده از هسته تابع پایه شعاعی

.(۲۰۲۲)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpc.html

طبقه‌بندی فرآیند گوسی

- در GPCتابع لجستیک به‌عنوان تابع پیوند معرفی می‌شود.
- توزیع احتمال پسین گوسی نیست.

- آموزش و پیش‌بینی در طبقه‌بندی فرآیند گاوسی به دو روش صورت می‌گیرد، یک در مقابل استراحت(One-versus-rest)و یک در مقابل یک(One-versus-one)
- ترکیب دو روش هم ممکن است.

۱. One-versus-rest برای جدا کردن یک کلاس از بقیه کلاس‌ها.

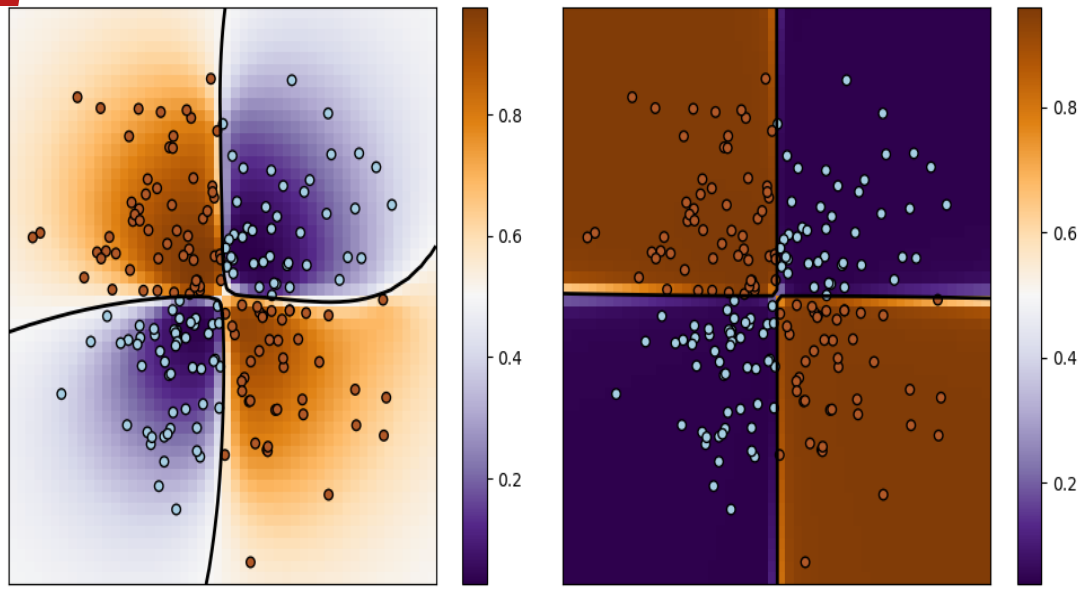
۲. One-versus-one برای جدا کردن دو کلاس از بقیه کلاس‌ها.

✓ One-versus-one ازنظر محاسباتی ارزان‌تر است زیرا مسائلی را حل می‌کند که فقط شامل یک زیرمجموعه از کل مجموعه آموزشی است.

مثال طبقه‌بندی فرآیند گاوسی

✓ طبقه‌بندی فرآیند گوسی روی مجموعه داده XOR

- هسته در فرآیند گوسی، به هسته ثابت (Isotropic) و هسته غیرساکن (Dot Product) تقسیم می‌شود.
- هسته ثابت مثل RBF



طبقه‌بندی فرآیند گاوسی در مجموعه داده XOR(هسته غیرساکن-هسته ثابت)

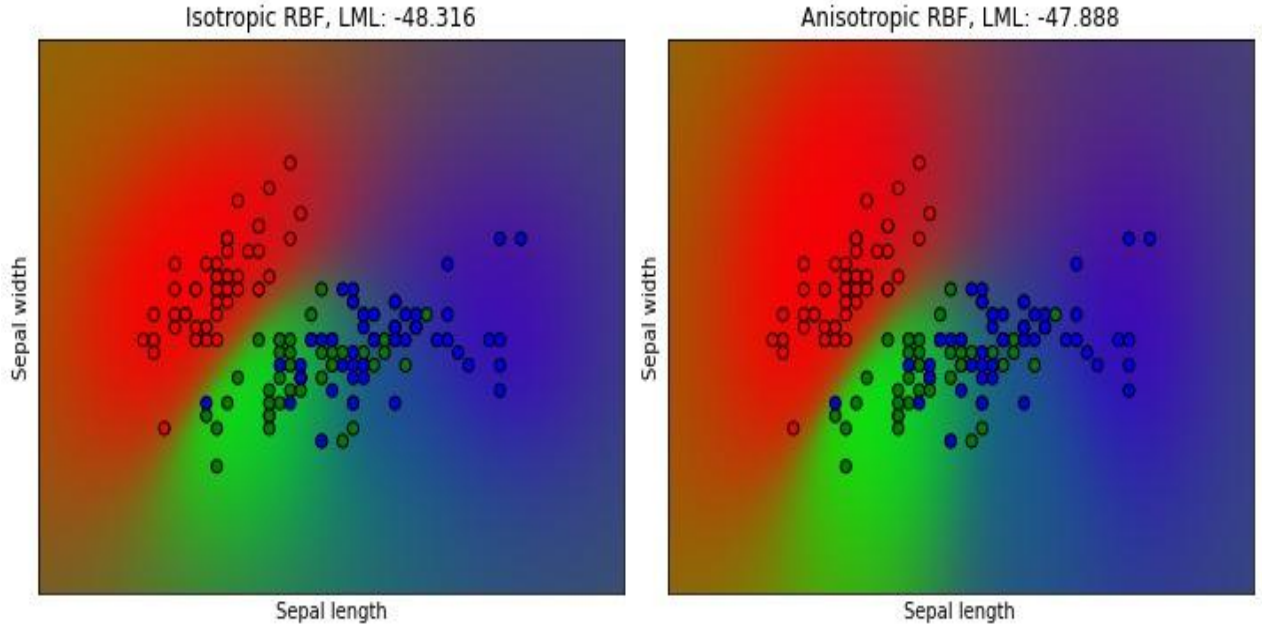
هسته غیرساکن نتایج بهتری به دست می‌آورد زیرا مرزهای کلاس خطی هستند.

✓ طبقه‌بندی فرآیند گاوسی در مجموعه داده عنبیه

.(۲۰۲۲/۰۱)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html

مثال طبقه‌بندی فرآیند گاوسی

- ✓ طبقه‌بندی فرآیند گاوسی در مجموعه داده عنبیه



همسانگرد و ناهمسانگرد RBFرا برای هسته GPCاحتمال پیش‌بینی‌شده

.(۲۰۲۲/۰۱)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html

هسته در فرآیندهای گاوسی

- هسته‌ها (که توابع کوواریانس نیز نامیده می‌شوند) یک جزء مهم فرآیندهای گوسی هستند که شکل احتمال پسین و احتمال پیشین فرآیند گوسی را تعیین می‌کنند.
- استفاده از دونقطه دیتا، با این فرض که نقاط داده مشابه، باید مقادیر هدف مشابهی داشته باشند.
- هسته‌های ثابت فقط به فاصله دونقطه داده وابسته هستند و به مقادیر مطلق آن‌ها بستگی ندارند.

- هسته ثابت به دو دسته تقسیم می‌شوند:

۱. همسانگرد

۲. ناهمسانگرد

- هسته‌ها توسط یک بردار از هایپرپارامترها نوشته می‌شوند. هایپرپارامترها برای مثال می‌توانند مقیاس طول یا تناوب یک هسته را کنترل کنند.

هسته در فرآیندهای گاوسی(ادامه)

- هسته ثابت (ConstantKernel) را می‌توان به‌عنوان نوعی از هسته نام برد که میانگین فرآیند گاوسی را تغییر می‌دهد و به پارامتر constant-value بستگی دارد.
- $k(x_i, x_j) = constant_value \forall x_1, x_2$
- $k(x_i, x_j)$ عملکرد هسته

❖ اپراتورهای هسته:

- هسته مجموع $k_{sum}(x, y) = k_1(x, y) + k_2(x, y)$
- هسته غیرثابت $k_{\mathrm{Product}}(x, y) = k_1(x, y) * k_2(x, y)$
- هسته توان $k_{exp}(x, y) = k(x, y)^p$

هسته در فرآیندهای گاوسی(ادامه)

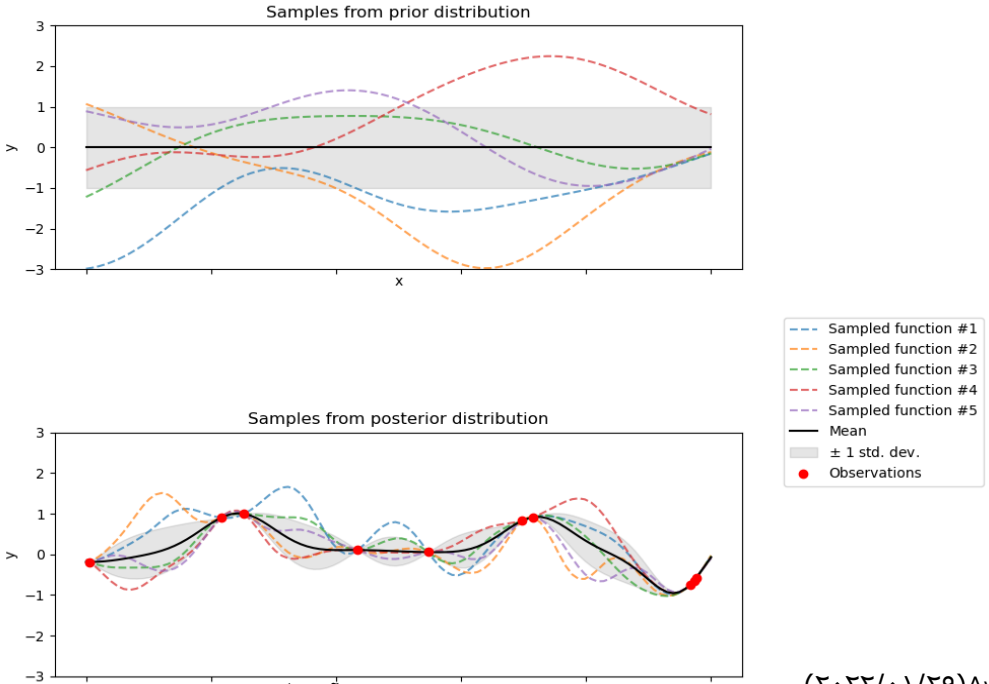
RBF

- هسته تابع پایه شعاعی(Radial-Basic Function)یک هسته ثابت است. و به آن هسته مربع نمایی(squared exponential) هم گفته می‌شود.

$$K(x_i, x_j) = exp\left(-\frac{d(x_i, x_j)^2}{2l^2}\right)$$

L یک اسکالر (نوع همسانگرد هسته) یا یک بردار برابر با تعداد ابعاد ورودیX (نوع ناهمسانگرد هسته) است.

$d(x_i, x_j)$ نشان دهنده فاصله اقلیدسی است.



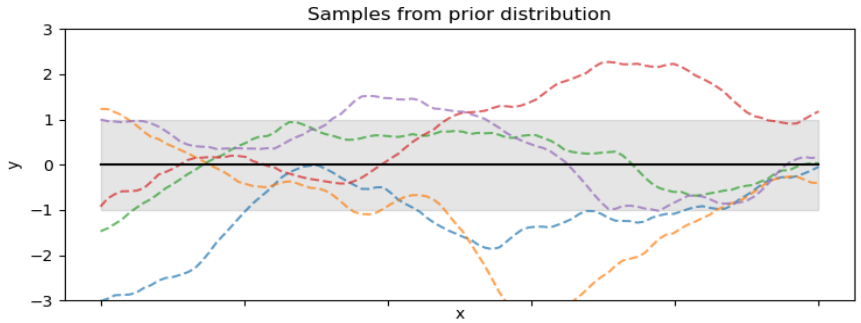
احتمال پیشین و احتمال پسین فرآیند گوسی(با هسته RBF)

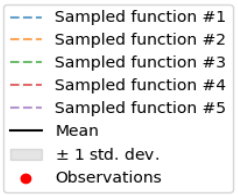
.(۲۰۲۲/۰۱/۲۹)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_prior_posterior.html

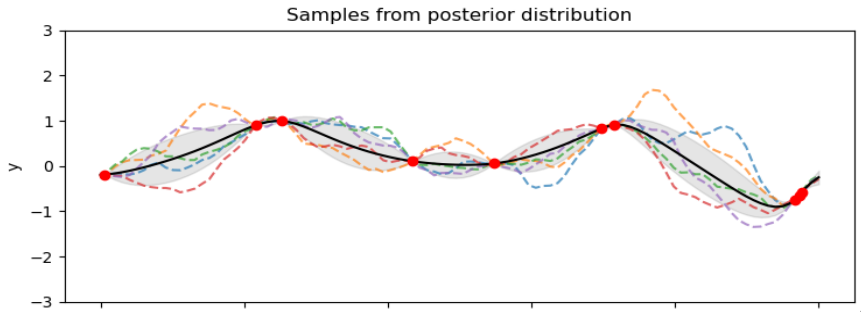
هسته در فرآیندهای گاوسی(ادامه)

- هسته ماترن (Matern kernel)یک هسته ثابت و تعمیم هسته RBF است.

$$K(x_i, x_j) = \frac{1}{\Gamma(v)2^{v-1}}\left(\frac{\sqrt{2v}}{l}d(x_i, x_j)\right)^{v} k_v\left(\frac{\sqrt{2v}}{l}d(x_i, x_j)\right)$$







احتمال پیشین و احتمال پسین فرآیند گوسی(با هسته ماترن)

L یک اسکالر (نوع همسانگرد هسته) یا یک بردار برابر با تعداد ابعاد ورودیX (نوع ناهمسانگرد هسته) است.

پارامتر vصاف و هموار بودن تابع حاصل را کنترل می‌کند.

dنشان دهنده فاصله اقلیدسی است.

k_v یک تابع بسل اصلاح‌شده و $\Gamma(v)$تابع گاما است.

وقتیv به سمت بی‌نهایت برود، هسته ماترن به هسته RBFهمگرا می‌شود.

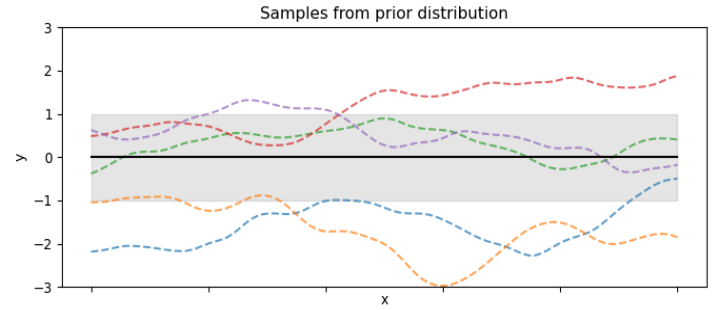
وقتی مطلق نمایی هسته با ماترن هسته $v = \frac{1}{2}$ یکسان می‌شود.

(۲۰۲۲/۰۱/۱۴)Available: https://scikit-learn.org/stable/modules/gaussian_process.html#duv2014.

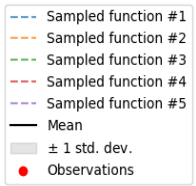
هسته در فرآیندهای گاوسی(ادامه)

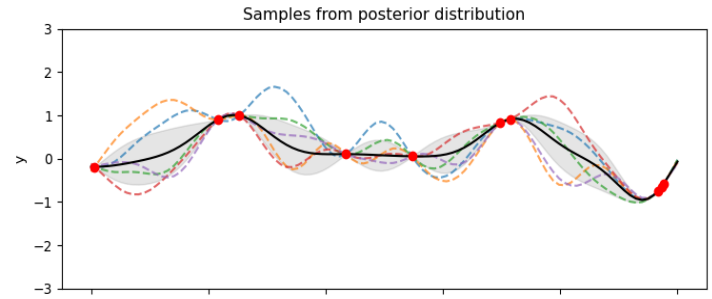
- ✓ هسته درجه دوم کسری(Rational quadratic kernel)را می‌توان به‌عنوان مجموع نامتناهی از هسته‌های RBFمشاهده کرد.

$$K(x_i, x_j) = \left(1 + \frac{d(x_i, x_j)^2}{2\alpha l^2}\right)^{-\alpha}$$



- ✓ هسته مربع سینوسی نمایی(ExpSineSquared)برای مدل‌سازی توابع پریودیک استفاده می‌شود.





$$K(x_i, x_j) = exp\left(\frac{2sin^2\left(\pi d(x_i, x_j)/p\right)}{2l^2}\right)$$

احتمال پیشین و پسین فرآیند گوسی با هسته درجهٔ دوم کسری

learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifie .(۲۰۲۲/۰۱/۰۵)Available: https://scikit-r

منابع

- [۱] .(۲۰۲۲) .*https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html*. Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html
- [۲] .(۲۰۲۲/۰۱)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html
- [۳] Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpr_noisy.html
- [۴] .(۲۰۲۲) .'Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_compare_gpr_krr.html
- [۵] .(۲۰۲۲)Available: https://scikit-learn.org/stable/auto_examples/gaussian_process/plot_gpc.html
- [۶] .(۲۰۲۲/۰۱/۲۰)Available: https://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier
- [۷] .(۲۰۲۲/۰۱/۰۵)Available: https://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier
- [۸] .(۲۰۲۲/۰۱/۲۶)Available: https//:scikit-learn.org/stable/modules/multiclass.html#multiclass

با تشکر از توجه شما