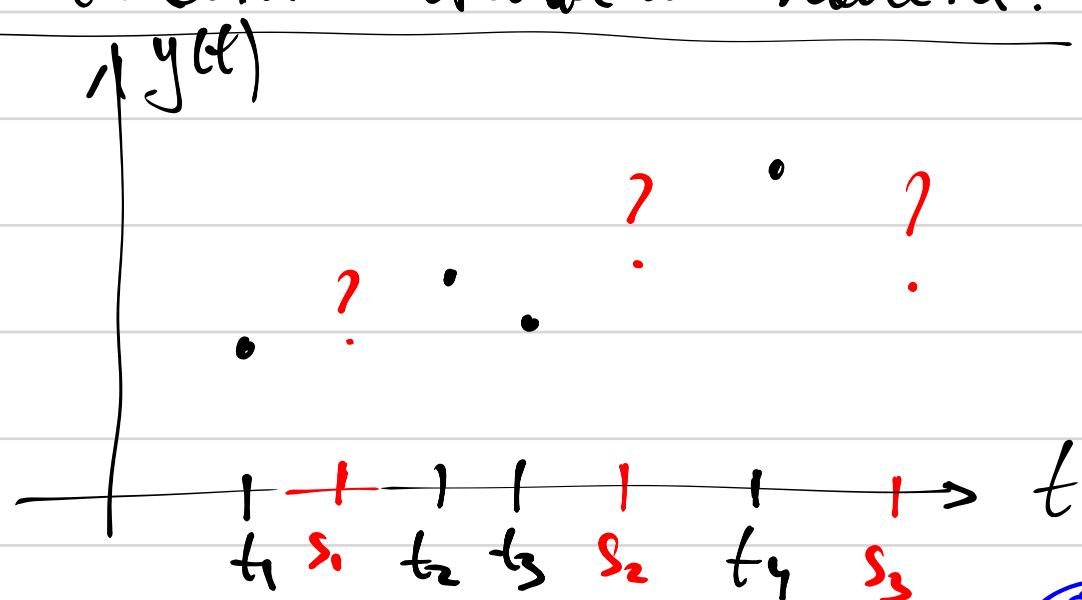
Видно / слышно?

→ Гауссовские процессы [Gaussian processes]

Зачем?

* врем. ряды
* оптимизация (если дорого считать значение ф-ции в точке)
* задача регрессии.

Вид с высоты птичьего полёта.



в GP всё норм-но!   надо опр-ть

② что можно?

вектор набл-й → y

вектор точек, что я хочу спрогн-ть (восстановить) → y^*

$$\begin{pmatrix} y \\ y^* \end{pmatrix} \sim N\left(\boxed{?};\ \boxed{?}\right)$$

① как перейти от $(y, y^*) \sim N$ к $(y^*|y) \sim ?$

$$(y^*|y) \sim N(\dots;\dots)$$

① нан от $\begin{pmatrix} a \\ b \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} ; \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \right)$

$$E\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

$$Var\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

упр. $(a|b) \sim ?$

Три подхода.

1-ый [классический с упором на линал].

теор вер — линал и блочные м-цы

$$\boxed{f(a|b) = \frac{f(a,b)}{f(b)}} = \dots \text{ упр- идаліся до ответа } =$$

$d^2(2\ln f(b))$

$$f(b) = \frac{1}{\sqrt{(2\pi)^{\dim b} \cdot \det \Sigma_{bb}}} \cdot \exp\left(-\frac{1}{2}(b-\mu_b)^T \Sigma_{bb}^{-1} (b-\mu_b)\right)$$

$f(a,b) = \dots\dots\dots$

$$= \frac{1}{\sqrt{(2\pi)^{\dim A} \det \Sigma_{a|b}}} \cdot \exp\left(-\frac{1}{2}(a-\mu_{a|b})^T \cdot \Sigma_{a|b}^{-1} \cdot (a-\mu_{a|b})\right)$$

линалом палим к такому виду и получаем ответ $\Sigma_{a|b} = \dots$

$\mu_{a|b} = \dots$

[этот подход первый учится]

2-ой подход.

[подход 1 + наблюд-сть!]

$$\begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}\right)$$

[идея 1] сначала рассмотрим:

$$\begin{pmatrix} a \\ b \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \Sigma\right)$$

$$f(a|b) = \frac{f(a,b)}{f(b)} = \frac{\text{конст}\cdot \exp(-a^T \cdot A\, a)}{\text{не зав от } a}$$

[идея 2] $(a|b) \sim N$

[идея 3] для многом норм-го:

незав-сть $\Leftrightarrow$ некорр-сть

часть a, к-рая не зав от b.

$$\boxed{a = \underbrace{M \cdot b}_{} + u}$$

на часть a, к-рая предсказ b

вектор	размер
a	$\sim [\dim a \times 1]$
b	$\sim [\dim b \times 1]$
M	$\sim [\dim a \times \dim b]$
u	$\sim [\dim a \times 1]$

подобрав M, добьёмся $Cov(b, u) = 0$

незав-сти b и u

$$E(a|b) = E(M\cdot b + u | b) = M\cdot b + E(u) = M\cdot b$$

$$Var(a|b) = Var(Mb + u | b) = Var(u|b) = Var(u)$$

$$\boxed{(a|b) \sim N(M\cdot b;\ Var(u))}$$

$Cov(b,u) = 0 \qquad Cov(b, a - Mb) = 0$

$$Cov(b,a) = Cov(b, Mb)$$

[переставим]

$$Cov(a,b) = Cov(Mb, b)$$

$$\Sigma_{ab} = Cov(a,b) = M \cdot Var(b) = M \cdot \Sigma_{bb}$$

$$M = \boxed{\Sigma_{ab} \cdot \Sigma_{bb}^{-1}} \quad !!$$

$$Var(a) = Var(Mb + u) = Var(Mb) + Var(u)$$

независ

$$Var(a) = M \cdot Var(b) \cdot M^T + Var(u)$$

$$\Sigma_{aa} = \Sigma_{ab} \cdot \underbrace{\Sigma_{bb}^{-1} \cdot \Sigma_{bb}}_{I} \cdot (\Sigma_{ab} \cdot \Sigma_{bb}^{-1})^T + Var(u)$$

$$\boxed{Var(u) = \Sigma_{aa} - \Sigma_{ab} \cdot \Sigma_{bb}^{-1} \cdot \Sigma_{ba}}$$

Итого для случая $\begin{pmatrix} a \\ b \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}\right)$

$$(a|b) \sim N(\Sigma_{ab} \cdot \Sigma_{bb}^{-1} b;\ \Sigma_{aa} - \Sigma_{ab} \cdot \Sigma_{bb}^{-1} \cdot \Sigma_{ba})$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}; \Sigma\right)$$

$$\boxed{(a|b) \sim N(\Sigma_{ab} \cdot \Sigma_{bb}^{-1} (b - \mu_b) + \mu_a;\ \Sigma_{aa} - \Sigma_{ab} \cdot \Sigma_{bb}^{-1} \cdot \Sigma_{ba})}$$

3-ий подход: теор. вер + оптимиз.
преобр. ф-ций.

! макс-ция нор-ой плотности по аргументу даёт мат. ожидание.

$$\boxed{f(a|b) \to \max_a} \qquad a^* = E(a|b)$$

аргумент плотности

случ.-ая величина

$$\boxed{f_{a|b}(t|s)}$$

$$f(a|b) = \frac{f(a,b)}{f(b)}$$

$$\ln f(a|b) = \ln f(a,b) - \ln f(b) \to \max_a$$

$$\ln f(a|b) = \ln\left[\frac{1}{\sqrt{(2\pi)^{\text{dimension}} \cdot \det \Sigma}} \cdot \exp\left(-\frac{1}{2}\begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix}^T \cdot \Sigma^{-1} \cdot \begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix}\right)\right] =$$

$$= \text{const} - \frac{1}{2}\begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix}^T \cdot \Sigma^{-1} \cdot \begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \qquad \Sigma^{-1} = V = \begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix}$$

$$\max_a \left[-\begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix}^T \cdot \begin{pmatrix} V_{aa} & V_{ab} \\ V_{ba} & V_{bb} \end{pmatrix} \cdot \begin{pmatrix} a-\mu_a \\ b-\mu_b \end{pmatrix} \right]$$

$$\max_a \underbrace{\left[-(a-\mu_a)^T \cdot V_{aa} \cdot (a-\mu_a) - 2 \cdot (a-\mu_a) \cdot V_{ab} (b-\mu_b) \right.}_{h(a)}$$

$$\max_a \underbrace{\left[-(a-\mu_a)^T \cdot V_{aa} \cdot (a-\mu_a) - 2\cdot(a-\mu_a)\cdot V_{ab}(b-\mu_b)\right]}_{h(a)}$$

$$dh(a) = -da^T \cdot \left[2\cdot V_{aa}\cdot(x_a-\mu_a) + 2\cdot V_{ab}(x_b-\mu_b)\right]$$

$$V_{aa}\cdot(x_a^*-\mu_a) + V_{ab}\cdot(x_b-\mu_b) = 0$$

$$x_a^* = E(a|b) = \mu_a - V_{aa}^{-1}\cdot V_{ab}\cdot(x_b-\mu_b)$$

[всп. линал ....] $\qquad V = \Sigma^{-1}$

[элегантно: еще немного вер-стей].

$$d^2h = -da^T \cdot V_{aa} \cdot da \cdot 2$$

$$\updownarrow$$

$$(\Sigma_{a|b})^{-1} \qquad \ldots$$

---

① какие предположения нужны про закон распр-ия.

во врем. рядах

$\longrightarrow t$

задача регрессии $\qquad \boxed{x_t = t}$

Модель:

Вр. ряд $[x_t = t]$

$$\begin{cases} y_t = f(x_t) + u_t \\ u_t \sim N(0; \sigma^2) \\ f \sim GP(m; k) \end{cases} \text{ незав.}$$

$$\boxed{f \sim GP(m; k)}$$

для любого набора $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$f(x) = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \sim N\left(m(x); k(x,x) \right)$$

$$\to \boxed{m(x)} = \begin{pmatrix} m(x_1) \\ m(x_2) \\ \vdots \\ m(x_n) \end{pmatrix}$$

$$\to \boxed{k(x,x)} = \begin{pmatrix} k(x_1,x_1) & k(x_1,x_2) & \dots \\ \vdots & & \\ k(x_1,x_n) & \dots & k(x_n,x_n) \end{pmatrix}$$

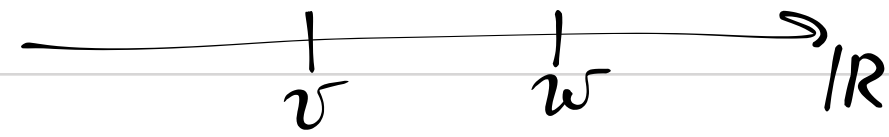
$\boxed{\text{Какие } m(x) \text{ и } k(x,x) \text{ можно исп-ть?}}$

$$m(x) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$m(x) = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} \to \text{пар-р и оцениваем с помощью Max Lik.}$$

$k(x,x) \leftarrow$ должна получ-ться неотр-но опр-ая м-ца для любого вектора x.

$K(v, w) =$



$= \langle \varphi(v), \varphi(w) \rangle$

Ядра в студию!

$$\begin{cases} RBF: & \alpha \cdot \exp(-\beta(v-w)^2) \\ LIN: & \alpha + \beta \cdot v \cdot w \\ COS: & \alpha \cdot \cos(\beta(v-w)) \end{cases}$$

[есть много ядер]

[есть искусство создавать новые]

$$K_{new} = K_1 + K_2$$
$$K_{new} = K_1 \cdot K_2$$
$$K_{new} = K_1(\varphi(v), \varphi(w))$$
$$K_{new} = \alpha \cdot K_1$$

ч. случай: $K_{new} = \langle \varphi(v), \varphi(w) \rangle$

PER: $\alpha \cdot \exp(-\beta \cdot \cos(\gamma(v-w)))$

$$LIN + PER + RBF_1 \cdot COS_1 + RBF_2 \cdot COS_2$$

искусство комбин-ния / мода