Гауссовские процессы

$$\begin{cases} y_t = f(x_t) + u_t \\ u_t \sim N(0; \sigma^2) \\ \boxed{f \sim GP(\mu; k)} \\ u_t \text{ и } f \text{ незав} \end{cases} \quad \forall t$$

$\boxed{\text{набл: } y_1, \dots, y_T}$

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_T) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x_1) \\ \vdots \\ \mu(x_T) \end{pmatrix}; \begin{bmatrix} k(x_1,x_1) & \dots & \\ \vdots & \ddots & \\ & & k(x_T,x_T) \end{bmatrix} \right)$$

для одномерного вр-го ряда: $\boxed{x_t = t}$ /в простой сл./

<u>Фаза 1</u> $\boxed{\text{Оценить пар-ры } \underline{\mu} \text{ и } \underline{k}}$ (метод макс. правд)

<u>Фаза 2</u>

$$\begin{matrix} \text{набл} \\ \\ \text{буд.} \end{matrix} \begin{pmatrix} y_1 \\ \vdots \\ y_T \\ \hline y_{T+1} \\ y_{T+2} \end{pmatrix} \sim N\left(\begin{pmatrix} \\ \\ \\ \end{pmatrix}; \begin{pmatrix} \\ \\ \\ \end{pmatrix} \right)$$

!!

$$\left(\begin{pmatrix} y_{T+1} \\ y_{T+2} \end{pmatrix} \middle| \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix} \right) \sim N(? ; ?)$$

Ещё одно приложение ГП
Байесовская оптимизация.

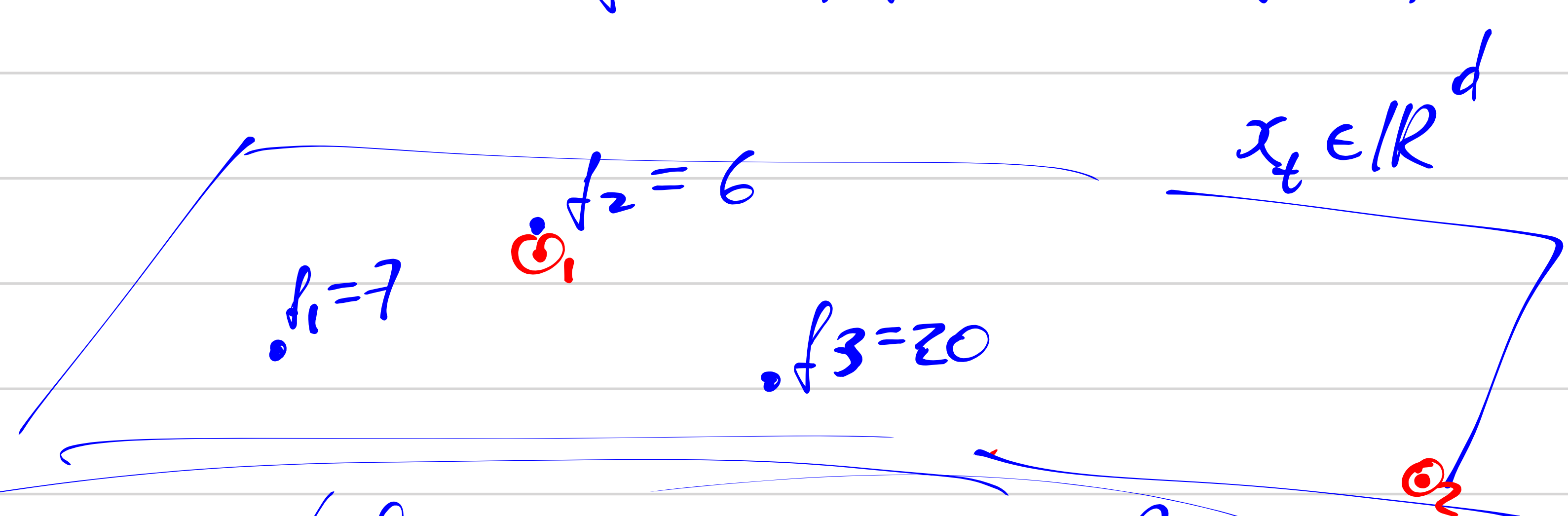
для опр-сти: min

ф-ция дорога для подсчёта!

Ф-ция явл-ся чёрным ящиком!

$$x=\begin{pmatrix}5\\3\\2\\6\end{pmatrix} \longrightarrow \square \longrightarrow f(x)=6$$

уже посчитали $f(x_1), f(x_2) \dots f(x_T)$



где попробовать точку x_{T+1}?

какие будут ф-ции μ и k?

$\mu(x)=0$ $\quad$ k – гаусс. ядро $\quad$ } совсем наив.

Идея 1 $\quad$ Эвристика $\quad$ Lower Confidence Bound!

инт: стоит пробовать точки с низким ожиданием и высокой дисперсией.

① $$\min_{x_{T+1}} E\big(f(x_{T+1}) \mid f_1, \dots f_T\big) - \gamma \cdot \sqrt{\mathrm{Var}\big(f(x_{T+1}) \mid f_1 \dots f_T\big)}$$

! сама f не считается !

② посчитали $f(x_{T+1})$

Идея 2. Вер-сть улучшения Probability of improvement

$$\max_{x_{T+1}} P\left(f(x_{T+1}) < y_{best} \,\middle|\, f_1 \dots f_T\right)$$

y_{best} – текущее достигнутое оптим-ое знач-ие

$$P\left(\frac{f(x_{T+1}) - E(f(x_{T+1}) | f_1 \dots f_T)}{\sqrt{Var(f(x_{T+1}) | f_1 \dots f_T)}} < \frac{y_{best} - E(f(x_{T+1}) | f_1 \dots f_T)}{\sqrt{Var(f(x_{T+1}) | f_1 \dots f_T)}}\right)$$

$$E(f(x_{T+1}) | f_1 \dots f_T) = m_{T+1}$$

$$Var(f(x_{T+1}) | f_1 \dots f_T) = \sigma^2_{T+1}$$

не требуют вычисления самой f, т.к. основано на предп и ф-ие условной вер-сти

$$\boxed{PI = \Phi\left(\frac{y_{best} - m_{T+1}}{\sigma_{T+1}}\right) \to \max_{x_{T+1}}}$$

! не требует вычисления f. !

- $\max PI$: находим x_{T+1}
- считаем $f(x_{T+1})$

Идея 3. Expected improvement

max ожидаемое улучшение

$$\max_{x_{T+1}} E\left[\max(y_{best} - f(x_{T+1}), 0) \,\middle|\, f_1, \dots f_T\right]$$

в этой вспом-ой задаче не исп-ся вычисление самой f, а исп-ся модель

$$E\left[\max\left(y_{best} - f(x_{T+1}), 0\right) \middle| f_1 \ldots f_T\right]$$

$$= \sigma_{T+1} \cdot E\left[\max\left(\underbrace{\boxed{\frac{y_{best} - m_{T+1}}{\sigma_{T+1}}}}_{= c} - \underbrace{\left(\frac{f(x_{T+1}) - m_{T+1}}{\sigma_{T+1}}\right)}_{N(0;1)}, 0\right) \middle| \ldots\right]$$

$$= \sigma_{T+1} \cdot E\left[\max\left(c - R, 0\right)\right]$$

$$R \sim N(0;1)$$

$$E\left[\max\left(c - R, 0\right)\right] =$$

плотность для $N(0;1)$

$$= \int_{-\infty}^{\infty} \max(c - z, 0) \cdot \phi(z)\, dz =$$

$$= \int_{-\infty}^{c} (c - z) \cdot \phi(z)\, dz = \boxed{\int_{-\infty}^{c} c \cdot \phi(z)\, dz} -$$

$$- \int_{-\infty}^{c} z\, \phi(z)\, dz = c \cdot \Phi(c) - \boxed{\int_{-\infty}^{c} z \cdot \phi(z)\, dz}$$

$$\int_{-\infty}^{c} z \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz =$$

$$= \frac{-1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \Bigg|_{z=-\infty}^{z=c} = -\phi(c)$$

Итого: $EI = c \cdot \Phi(c) + \phi(c)$ $\qquad c = \dfrac{y_{best} - m_{T+1}}{\sigma_{T+1}}$

новую точку x_{T+1} находим
из оптимизации

$$\max_{x_{T+1}} EI$$

считаем $f(x_{T+1})$

Векторная авторегрессия

$y_t \sim ARMA(p,q)$
$y_t \in \mathbb{R}$

можно обобщить !
$y_t \in \mathbb{R}^d$

→ довольно популярн.
→ громоздкие / много пар-в
→ много вариаций
→ куча комбинаций : стац/нестац

предп. $y_t \in \mathbb{R}^d$ – стац- вес многом. пос.

Опр.

① $E(y_t) = \mu \in \mathbb{R}^d$ [не зав-т от времени]

② $Cov(y_t, y_{t-k}) = \gamma_k \in \mathbb{R}^{d\times d}$ [не зав-т от врем]

для 2-го

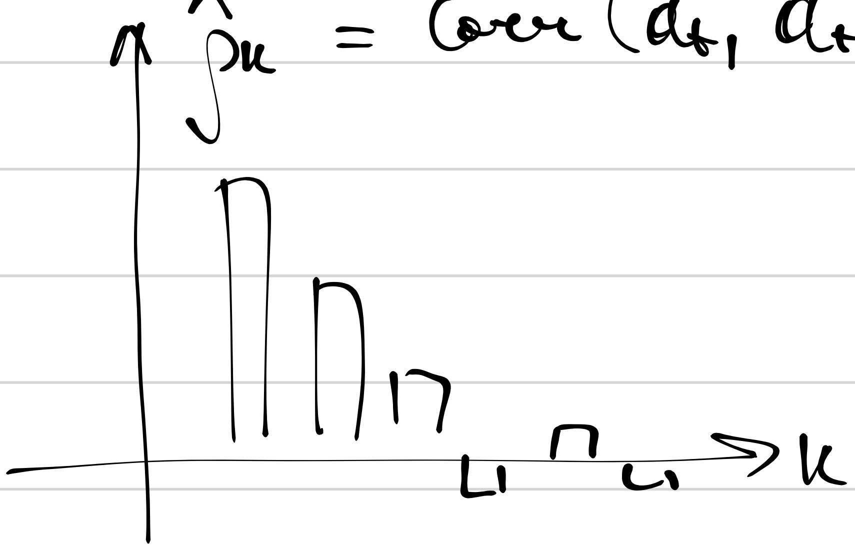
$$y_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix}$$

не зависит от времени

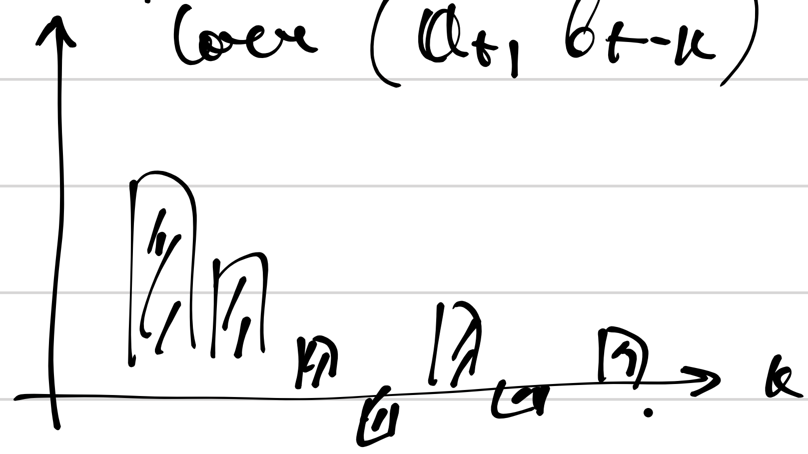
$E(a_t) = \mu_a \quad Cov(a_t, a_{t-k}) = \gamma_k^{aa}$
$E(b_t) = \mu_b \quad Cov(b_t, b_{t-k}) = \gamma_k^{bb}$
$Cov(a_t, b_{t-k}) = \gamma_k^{ab}$

в кор-м анализе.

$$\hat\rho_k = \widehat{Corr}(a_t, a_{t-k})$$



доп-но появляется
кросс-корр-ция
$\widehat{Corr}(a_t, b_{t-k})$



MA, AR, ARMA → VMA, VAR, VARMA.

опр $\boxed{\text{VMA}}$ = vector moving average.
VMA(q)

[упр] Сколько пар-в оц-ся?
$d + q \cdot d \times d + \left(d + \frac{d \cdot (d-1)}{2}\right)$

$$\boxed{y_t = \mu + u_t + A_1 \cdot u_{t-1} + \ldots + A_q \cdot u_{t-q}}$$

$y_t \in \mathbb{R}^d$ $\quad \mu \in \mathbb{R}^d$

$A_1, A_2 \ldots A_q \in \mathbb{R}^{d \times d}$

опр. $\boxed{u_t \text{ – многомерный белый шум.}}$

$u_t \in \mathbb{R}^d$ $\quad E(u_t) = 0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

$Cov(u_t, u_{t-k}) = 0 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ при $k \neq 0$

$Var(u_t) = \Sigma \in \mathbb{R}^{d \times d}$

Теор. По построению $\overset{\text{любой}}{\text{ИМА}(q)}$ стац-н.