Лекция 3.
Основы теории статистических решений

Оценивание – принятие решения о значении сообщения в каждый момент, либо в заданные моменты времени

Основной вопрос теории статистического синтеза:

«Какая система оценивает сообщения наилучшим образом ???»

Определения понятий

$$y(t) = S(t, \lambda(t)) + n(t)$$

Наблюдение на интервале времени $[0,t] \rightarrow$ реализация Y_0^t

Сигнал

Сообщение $\lambda(t) = \mathbf{c}\mathbf{x}(t)$

$\mathbf{x}(t) -$ векторный марковский процесс с априорной ПВ: $p(\mathbf{x},t)$

Шум/помеха

Апостериорная ПВ: $p(\mathbf{x},t \mid Y_0^t)$

Определение понятий

<u>Решение</u> – результат обработки информации в РТС

<u>Принятие решения</u> – процесс обработки информации в РТС

<u>Решающее правило</u> – алгоритм обработки (u)

<u>Функция потерь</u> (c) – показатель качества алгоритма обработки

$\mathbf{d} = \mathbf{u}\left(Y_0^t\right)$; $\mathbf{d}$- результат решения

$\mathbf{u}\left(*\right)$ - решающее правило

$c\left(\mathbf{x},\mathbf{d}\right)$ - функция потерь

Функция потерь

Свойства:

1) $c(\mathbf{x},\mathbf{d})$ - скалярная

2) $c(\mathbf{x},\mathbf{x})=0$ (потери равны нулю, если $\mathbf{d}=\mathbf{x}$)

3) $c(\mathbf{x},\mathbf{d}_1)\geq c(\mathbf{x},\mathbf{d}_2)$ всегда, если $\|\mathbf{x}-\mathbf{d}_1\|\geq\|\mathbf{x}-\mathbf{d}_2\|$

4) $c(\mathbf{x},\mathbf{d}_1)=c(\mathbf{d}_1,\mathbf{x})$

Примеры:

1) квадратичная $c(\mathbf{x},\mathbf{d})=\|\mathbf{x}-\mathbf{d}\|_Q^2$,

2) модульная $c(\mathbf{x},\mathbf{d})=\|\mathbf{x}-\mathbf{d}\|_Q$,

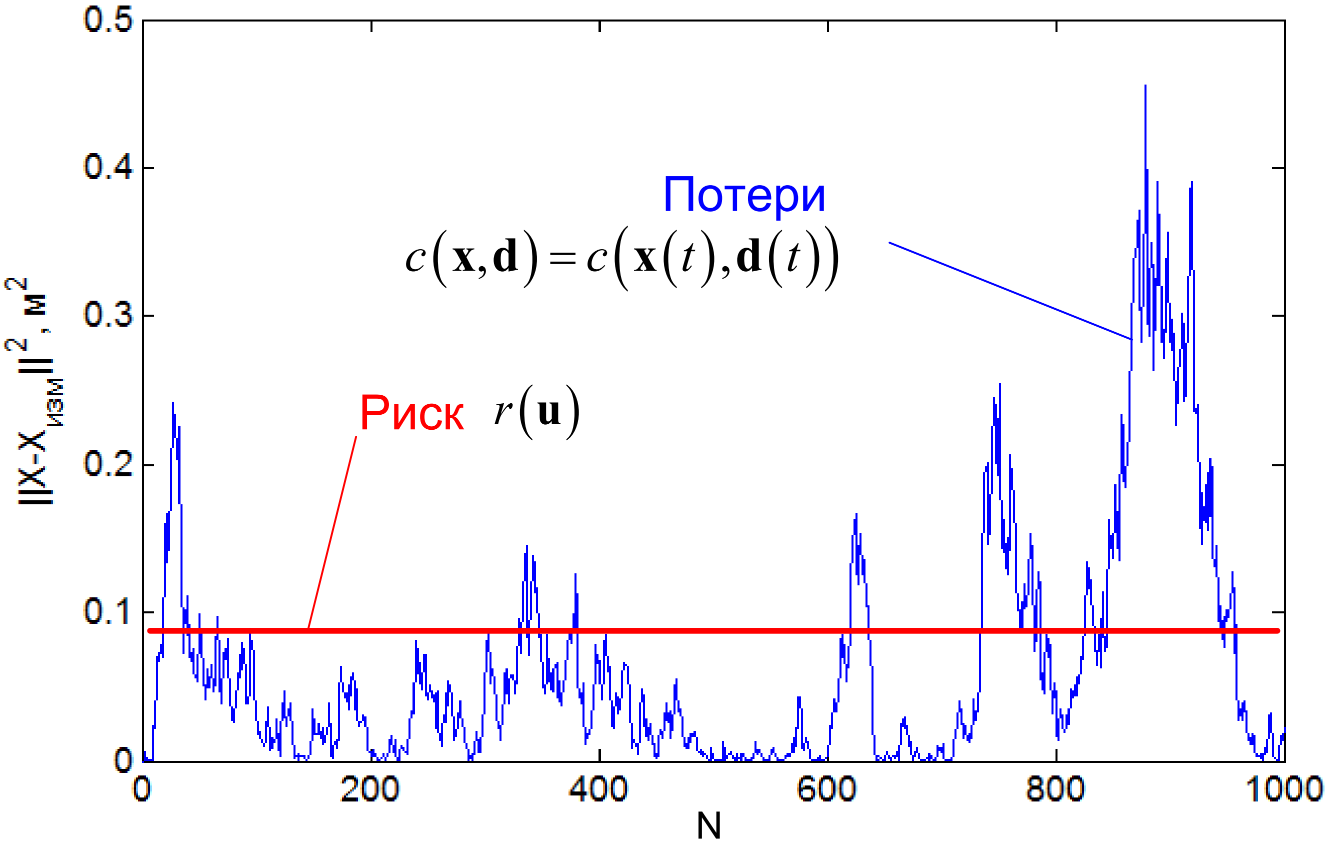
3) простая $c(\mathbf{x},\mathbf{d})=\begin{cases} 0, & \|\mathbf{x}-\mathbf{d}\|_Q<\varepsilon/2, \\ 1/\varepsilon, & \|\mathbf{x}-\mathbf{d}\|_Q\geq\varepsilon/2, \end{cases}$.

Здесь $\|*\|$ – символ нормы вектора, $\varepsilon=const\to 0$

Риск

Потери меняются во времени: $c(\mathbf{x},\mathbf{d}) = c(\mathbf{x}(t),\mathbf{d}(t))$

<u>Риск</u> – усредненные потери.



Средний риск: $$r(\mathbf{u},t) = \iint c\left(\mathbf{x}(t),\mathbf{u}\left(Y_0^t\right)\right) p\left(\mathbf{x}(t),Y_0^t\right) d\mathbf{x}(t)\, dY_0^t$$

Оптимальные решения

Байесовские решения – минимизируют средний риск

$$\mathbf{u}_0 = \arg\min_{\mathbf{u}} r(\mathbf{u}) = \arg\min_{\mathbf{u}} \iint c\left(\mathbf{x}, \mathbf{u}\left(Y_0^t\right)\right) p\left(\mathbf{x}, Y_0^t\right) d\mathbf{x} dY_0^t$$

Оптимальное решающее правило для квадратичной функции потерь – апостериорное среднее:

$$\mathbf{u}_0\left(Y_0^t\right) = \hat{\mathbf{x}} = \int \mathbf{x} p\left(\mathbf{x} \middle| Y_0^t\right) d\mathbf{x}$$

Оптимальное решающее правило для простой функции потерь – максимум апостериорной ПВ:

$$\mathbf{u}_0\left(Y_0^t\right) = \hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p\left(\mathbf{x} \middle| Y_0^t\right)$$

Оптимальные решения

Минимаксные решения – минимизируют максимум условного риска

$$\tilde{r}(\mathbf{x},\mathbf{u}) = \int c(\mathbf{x},\mathbf{u}(Y))p(Y|\mathbf{x})dY$$

Оптимальное решающее правило:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \max_{\mathbf{x}} \tilde{r}(\mathbf{x},\mathbf{u})$$

Метод максимального правдоподобия

Функция правдоподобия, это условная ПВ рассматриваемая как функция **x**: $L(\mathbf{x}) = p(Y(\mathbf{x}_{\text{и}}) \,|\, \mathbf{x})$

Оценка максимального правдоподобия: $$\left.\frac{\partial L(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}_{\text{м}}} = 0$$

Учёт наличия случайных неинформативных параметров сигнала или сообщения

$$y(t) = S\left(t, \lambda(t), \boldsymbol{\mu}\right) + n(t)$$

$\boldsymbol{\mu}$ - вектор (набор) случайных неинформативных параметров

Необходимо усреднять средний риск по неинформативным параметрам:

$$r(\mathbf{u}) = \iiint c\left(\mathbf{x}, \mathbf{u}(Y, \boldsymbol{\mu})\right) p\left(\mathbf{x}, \boldsymbol{\mu}, Y\right) d\mathbf{x} d\boldsymbol{\mu} dY$$

Функция правдоподобия при наличии неинформативных параметров:

$$\widetilde{L}(\mathbf{x}) = \int p\left(Y \middle| \mathbf{x}, \mu\right) p(\mu) d\mu$$