

Лекция 2.

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: $p(\mathbf{x}, t | Y_0^t)$

Определение понятий

Решение – результат обработки информации в РТС

Принятие решения – процесс обработки информации в РТС

Решающее правило – алгоритм обработки (u)

Функция потерь (c) – показатель качества алгоритма обработки

$\hat{\mathbf{x}} = \mathbf{u}(\mathbf{Y}_0^t)$; $\hat{\mathbf{x}}$ - результат решения (оценка)

$\mathbf{u}(\ast)$ - решающее правило

$c(\mathbf{x}, \hat{\mathbf{x}})$ - функция потерь

Функция потерь

- Свойства:
- 1) $c(\mathbf{x}, \hat{\mathbf{x}})$ - скалярная
 - 2) $c(\mathbf{x}, \mathbf{x})=0$ (потери равны нулю, если $\hat{\mathbf{x}} = \mathbf{x}$)
 - 3) $c(\mathbf{x}, \hat{\mathbf{x}}_1) \geq c(\mathbf{x}, \hat{\mathbf{x}}_2)$ всегда, если $\|\mathbf{x} - \hat{\mathbf{x}}_1\| \geq \|\mathbf{x} - \hat{\mathbf{x}}_2\|$
 - 4) $c(\mathbf{x}, \hat{\mathbf{x}}_1) = c(\hat{\mathbf{x}}_1, \mathbf{x})$

Примеры:

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

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

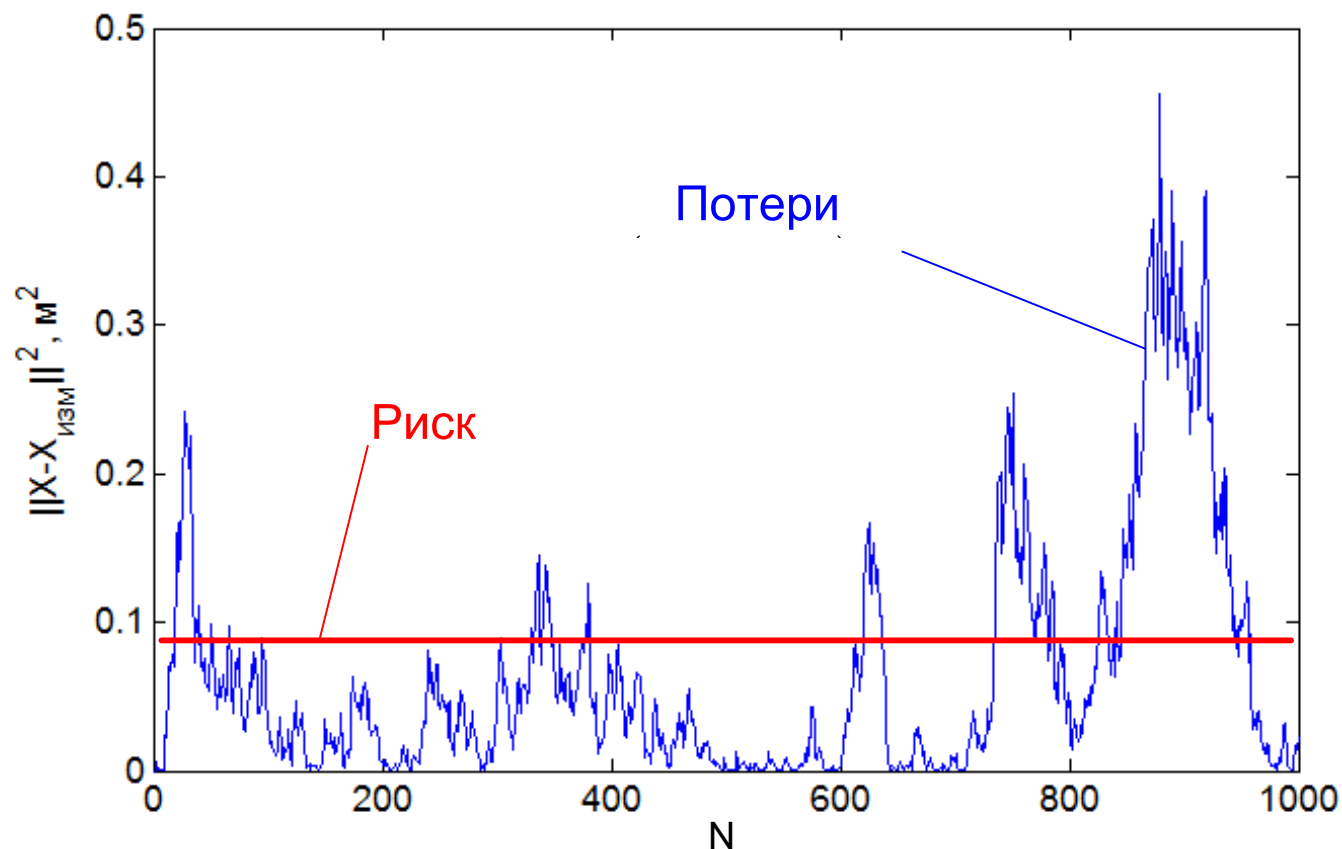
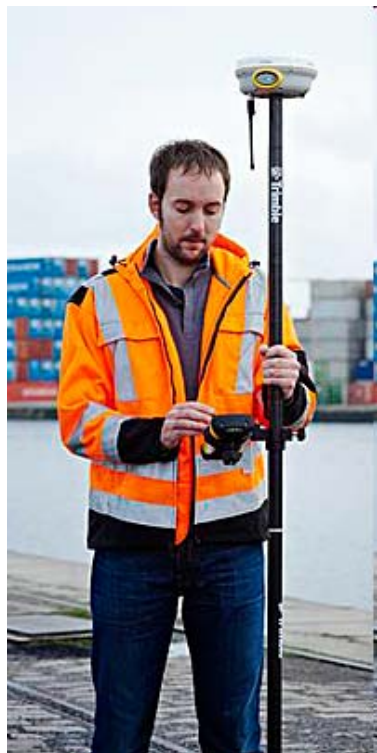
3) простая $c(\mathbf{x}, \hat{\mathbf{x}}) = \begin{cases} 0 & \text{при } \|\mathbf{x} - \hat{\mathbf{x}}\| < \varepsilon / 2, \\ 1/\varepsilon, & \|\mathbf{x} - \hat{\mathbf{x}}\| \geq \varepsilon / 2. \end{cases}$

Здесь $\|*\|$ – символ нормы вектора, $\varepsilon = \text{const} \rightarrow 0$

Риск

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

Риск – усредненные потери.



Средний риск:

$$r(\hat{\mathbf{x}}, t) = \iint c(\mathbf{x}(t), \hat{\mathbf{x}} = \mathbf{u}(\mathbf{Y}_0^t)) p(\mathbf{x}(t), \mathbf{Y}_0^t) d\mathbf{x}(t) d\mathbf{Y}_0^t$$

Оптимальные решения

Байесовские решения – минимизируют средний риск

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

Оптимальное решающее правило для квадратичной функции потерь – апостериорное среднее:

$$\mathbf{u}_0(\mathbf{Y}_0^t) = \hat{\mathbf{x}} = \int \mathbf{x} p(\mathbf{x} | \mathbf{Y}_0^t) d\mathbf{x}$$

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

$$\mathbf{u}_0(\mathbf{Y}_0^t) = \hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{x} | \mathbf{Y}_0^t)$$

Оптимальные решения

Минимаксные решения – минимизируют максимум условного риска

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

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 \mathbf{x} : $L(\mathbf{x}) = p(\mathbf{Y}(\mathbf{x}_и) | \mathbf{x})$

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

Учёт наличия случайных неинформативных параметров сигнала или сообщения

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

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

Необходимо усреднять средний риск по неинформативным параметрам:

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

Функция правдоподобия при наличии неинформативных параметров:

$$\tilde{L}(\mathbf{x}) = \int p(\mathbf{Y} | \mathbf{x}, \boldsymbol{\mu}) p(\boldsymbol{\mu}) d\boldsymbol{\mu}$$