Анализ экономических данных: машинное обучение

Дмитрий В. Виноградов

26 июня 2008 г.

МИНИМИЗАЦИЯ СРЕДНЕГО РИСКА

Определение 1 Пусть на множестве *примеров* X задан класс $C = \{c : X \to Y\}$ возможных *зависимостей* и определен функционал *среднего риска - критерий качества* выбранной зависимости

$$I(c) = \int_X \Phi(x, c(x)) dP(x), \qquad (1)$$

где $\int \ldots dP(x)$ - некоторый интеграл на множестве X. Функция $\Phi : X \times Y \to R^1$ называется *функцией потерь*. Требуется найти такую зависимость $h \in C$, чтобы значение $I(h)$ было минимальным.

Пример 1 Если $X = R^1$ и $\int \ldots dP(x) = \int \ldots p(x) dx$ - интеграл Римана относительно плотности $p(x)$, то получаем задачу *вариационного исчисления*.

ВОССТАНОВЛЕНИЕ ЗАВИСИМОСТЕЙ

Определение 2 Если мера $P(A) = \int_A dP(x)$ на множестве X неизвестна, но задана конечная *обучающая выборка* $S = \{(x_j, c(x_j)) : j = 1, \ldots, m\}$ для фиксированной неизвестной зависимости $c \in C$, примеры $x_1, \ldots, x_m \in X$ которой являются независимыми исходами с распределением P, то минимизация функции потерь Φ по мере P называется задачей *восстановления зависимости по эмпирическим данным*.

Утверждение 1 Пусть $X \subseteq [a, b] \subset R^1$, $\Phi(x, c(x)) = (p(x) - c(x))^2$ для некоторого неизвестного полинома $p(x)$ степени n. Тогда для полинома Лагранжа $l(x) = \sum_{(z,p(z)) \in S} p(z) \frac{\prod_{t \in \pi_1 S \setminus \{z\}} (x-t)}{\prod_{t \in \pi_1 S \setminus \{z\}} (z-t)}$ имеем $P[I(l) < \min\{I(p) : p \in C\} + \varepsilon] \geq 1 - \delta$ для достаточно большого объема $m(\varepsilon, \delta, n)$ выборки S, где C - класс полиномов.

РАСПОЗНАВАНИЕ ОБРАЗОВ

Определение 3 Пусть из множества X независимо извлекаются примеры $x_1, \ldots, x_m$ согласно неизвестного р.в. P и учитель классифицирует их согласно неизвестной функции *условного распределения* $P(k|x)$, $k \in \{0, 1\}$. Функционал среднего риска задается правилом:

$$I(c) = \Sigma_{k \in \{0,1\}} \int (k - c(x))^2 P(k|x) dP(x). \tag{2}$$

По выборке $S = (x_1, c(x_1)), \ldots, (x_m, c(x_m))$ требуется найти такую зависимость $h : X \to \{0, 1\}$, $h \in C$, чтобы для произвольно малой *ошибки* $\varepsilon \geq 0$ и произвольно малого $\delta \geq 0$ для $m \geq m(\varepsilon, \delta)$ выполнялось $P[I(h) < \min\{I(c) : c \in C\} + \varepsilon] \geq 1 - \delta$.

РЕГРЕССИЯ

Определение 4 Пусть имеется стохастическая зависимость между элементами множества X и элементами множества Y, задаваемая функцией *условного распределения* $P(y|x)$, $x \in X$. *Регрессия* y на x - это функция: $y(x) = \int yP(y|x)dy$.

Задача 1 *Доказать, что для регрессии* $y(x)$ *имеем*

$$\int (y - y(x))P(y|x)dy = 0. \tag{3}$$

ВОССТАНОВЛЕНИЕ РЕГРЕССИИ

Определение 5 Пусть из множества X независимо извлекаются примеры $x_1, \ldots, x_m$ согласно неизвестного р.в. P и согласно условным распределениям $P(y|x_j)$ случайным образом выбираются y_j. По выборке $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ требуется найти такую зависимость $h : X \to Y$, чтобы значение

$$I(h) = \iint (y - h(x))^2 P(y|x) dy dP(x). \tag{4}$$

было минимальным. Эта задача называется задачей *восстановления регрессии*.

Задача 2 *Доказать, что*

$$\begin{aligned}\iint (y - h(x))^2 P(y|x) dy dP(x) = \iint (y - y(x))^2 P(y|x) dy dP(x) \\ + \iint (y(x) - h(x))^2 P(y|x) dy dP(x)\end{aligned} \tag{5}$$

СЛУЧАЙНЫЕ ФУНКЦИИ

Определение 6 *Случайной функцией* со значениями в измеримом пространстве $(E, \mathcal{B})$ называют такую функцию $X : T \times \Omega \to E$, что для каждого $t \in T$ $X_t = X(t, \cdot) : \Omega \to E$ является измеримым отображением $(\Omega, \mathcal{F})$ в $(E, \mathcal{B})$, т.е. $\forall A [A \in \mathcal{B} \Rightarrow X_t^{-1}(A) \in \mathcal{F}]$.

Определение 7 *Реализацией* с.ф. $X : T \times \Omega \to E$ называется функция $X(\cdot, \omega) : T \to E$ при фиксированном $\omega \in \Omega$. Обычно реализация обозначается через $x(t)$.

Определение 8 *Случайным процессом* с фазовым пространством $(E, \mathcal{B})$ называют с.ф. $X : T \times \Omega \to E$, где $T = R^1$, или $T = [a, b] \subset R^1$.

СЛУЧАЙНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

Определение 9 *Случайной последовательностью* элементов измеримого пространства $(E, \mathcal{B})$ называют с.ф. $X : T \times \Omega \to E$, где $T = Z$, или $T = N$, или $T = \{1, \ldots, m\}$.

Пример 2 Если $X_1, \ldots, X_m$ - набор действительных с.в., то функция $X : \{1, \ldots, m\} \times \Omega \to R^1$, определенная правилом $X(n, \omega) = X_n(\omega)$, является случайной последовательностью.

Задача 3 *Какое свойство R^m позволяет утверждать, что набор с.в. $X_1, \ldots, X_m : \Omega \to R^1$ задает случайный вектор $\vec{X} : \Omega \to R^m$ - реализацию с.п. из предыдущего примера?*

ЦЕПИ МАРКОВА

Рассмотрим счетное фазовое пространство E, которое отождествим с N или с $\{0, \dots, r\}$. Пусть $\mathcal{B}$ - булеан всех подмножеств E. Это - дискретное пространство с метрикой $\rho(i, j) = 1$ при $i \neq j$ и $\rho(j, j) = 0$.

Определение 10 *Цепь Маркова* - это такая с.ф. $X : T \times \Omega \to E$, что для любого $m \geq 1$, всех точек $s_1 < \dots < s_m < s \leq t$ и любых $i, j, i_1, \dots, i_m \in E$ выполнено

$$P(X_t = j | X_{s_1} = i_1, \dots, X_{s_m} = i_m, X_s = i) = P(X_t = j | X_s = i). \quad (6)$$

ПЕРЕХОДНЫЕ ВЕРОЯТНОСТИ

Определение 11 *Переходными вероятностями* цепи Маркова называются функции $p_{ij}(s,t) = P(X_t = j | X_s = i)$, где $s \le t, (s,t \in T)$ и любых $i, j \in E$.

Задача 4 *Проверить, что*
$p_{ij}(s,t) \ge 0$ *для любых* $s \le t$ *и любых* $i, j \in E$.
$\sum\limits_{j \in E} p_{ij}(s,t) = 1$ *для любых* $s \le t$ *и любого* $i \in E$.
$p_{ij}(s,s) = \delta_{ij}$ *для любых* $s \in T$ *и любых* $i, j \in E$

Задача 5 (уравнение Маркова-Колмогорова) *Доказать, что для* $s \le u \le t$ *и любых* $i, j \in E$

$$p_{ij}(s,t) = \sum_{k \in E} p_{ik}(s,u) p_{kj}(u,t) \qquad (7)$$

КОНЕЧНОМЕРНЫЕ РАСПРЕДЕЛЕНИЯ

Определение 12 Для любых $t_1, \ldots, t_m \in T$ распределение с.в. $(X(t_1, \cdot), \ldots, X(t_m, \cdot) : \Omega \to R^m$ называется*конечномерным распределением* случайной функции $X : T \times \Omega \to E$.

Определение 13 *Начальным распределением* цепи Маркова называется распределение с.в.$X(0, \cdot) : \Omega \to R^1$. Это - набор неотрицательных чисел $p_i(0) = P(X_0 = i)$, в сумме дающих 1.

Теорема 2 Конечномерные распределения цепи Маркова однозначно определяются начальным распределением и переходными вероятностями:

$$P[(X_{t_1}, \ldots, X_{t_m}) \in B] = \\ = \sum_{(j_1,\ldots,j_m)\in B} \sum_i p_i(0) p_{ij_1}(0, t_1) \cdots p_{j_{m-1}j_m}(t_{m-1}, t_m). \quad (8)$$

ОДНОРОДНЫЕ ЦЕПИ МАРКОВА

Определение 14 Цепь Маркова называется *однородной*, если для всех $t \in T$ выполнено $p_{ij}(t, t+s) = p_{ij}(0, s) = p_{ij}(s)$.

Задача 6 *Доказать, что для* $s, t \in T$ *и любых* $i, j \in E$

$$p_{ij}(s+t) = \sum_{k \in E} p_{ik}(s) p_{kj}(t) \tag{9}$$

Теорема 3 (эргодическая) Пусть для однородной цепи Маркова $X = \{X_t, t \geq 0\}$ существуют такие $j_0 \in E$, $h > 0$ и $0 < \delta \leq 1$, что $p_{ij_0}(h) \geq \delta$ при всех $i \in E$.
Тогда для любого $j \in E$ существует независящий от начального состояния $i \in E$ предел $\lim\limits_{t \to \infty} p_{ij}(t) = \tilde{p}_j$, причем для всех $t > 0$

$$\left| p_{ij}(t) - \tilde{p}_j \right| \leq (1 - \delta)^{[t/h]} \tag{10}$$

СЛЕДСТВИЯ ЭРГОДИЧЕСКОЙ ТЕОРЕМЫ

Следствие 1 В условиях эргодической теоремы для любого $j \in E$ имеем $\lim\limits_{t\to\infty} p_j(t) = \tilde{p}_j$, где $p_j(t) = P(X_t = j)$ с оценкой скорости сходимости

$$\left|p_j(t) - \tilde{p}_j\right| \leq (1-\delta)^{[t/h]} \tag{11}$$

Следствие 2 В условиях эргодической теоремы для любых $t \in T$ и $j \in E$ имеем

$$\tilde{p}_j = \sum_i \tilde{p}_i p_{ij}(t). \tag{12}$$

Следствие 3 В условиях эргодической теоремы либо $\sum\limits_j \tilde{p}_j = 1$, либо все $p_j = 0$.

СТАЦИОНАРНЫЕ ПРОЦЕССЫ

Определение 15 Случайный процесс X называется *стационарным* в строгом смысле, если для любых $h, t_1, \dots, t_m \in T$ верно $P[(X_{t_1}, \dots, X_{t_m}) \in B] = P[(X_{t_1+h}, \dots, X_{t_m+h}) \in B]$.

Теорема 4 Пусть однородная цепь Маркова $X = \{X_t, t \geq 0\}$ с переходными вероятностями $p_{ij}(s,t) = P(X_t = j | X_s = i)$ имеет стационарное р.в. $\tilde{p}_j : j \in E$.
Тогда цепь Маркова $Y = \{Y_t, t \geq 0\}$ с $P(Y_t = j | Y_s = i) = p_{ij}(s,t)$ и начальным распределением $\tilde{p}_j : j \in E$ является стационарным процессом.

ВЕРОЯТНОСТНЫЕ АВТОМАТЫ

Определение 16 *Вероятностный автомат* (ВА) M - это объект $(Q, \Sigma, \tau, \gamma, \pi)$, где
Q - конечное множество *состояний*, Σ - конечный *алфавит*,
$\tau : Q \times \Sigma \to Q$ - функция (детерминированных) *переходов*,
$\pi : Q \to [0, 1]$ - начальное распределение: $\sum_{q \in Q} \pi(q) = 1$,
$\gamma : Q \times \Sigma \to [0, 1]$ - распределение следующего символа в текущем состоянии: $\sum_{\sigma \in \Sigma} \gamma(q, \sigma) = 1$ для каждого $q \in Q$.

Задача 7 *Доказать, что вероятность того, что M породит строку $\sigma_1, \ldots, \sigma_l$ равна:*

$$P_M(\sigma_1, \ldots, \sigma_l) = \sum_{q^0 \in Q} \pi(q^0) \prod_{i=1}^{l} \gamma(q^{i-1}, \sigma_i), \tag{13}$$

где $q^i = \tau(q^{i-1}, \sigma_i)$

СУФФИКСНЫЕ ВЕРОЯТНОСТНЫЕ АВТОМАТЫ

Определение 17 *Суффиксами* строки $s = \sigma_1, \ldots, \sigma_l$ называются строки $\sigma_k, \ldots, \sigma_l$ для $1 \leq k \leq l$ и пустая строка λ. Множество всех суффиксов строки s обозначается $Suffix^*(s)$. Множество S строк называется *безсуффиксным*, если $\forall s \in S$, $Suffix^*(s) \cap S = \{s\}$.

Определение 18 *Суффиксный* ВА M - это такой ВА, что $Q \subset \Sigma^*$ - безсуффиксное множество строк, и $\tau(s, \sigma)$ - суффикс $s \cdot \sigma$ для любого состояния $s \in Q$ и любого символа $\sigma \in \Sigma$.

Пример 3 Пусть $\Sigma = \{0, 1\}$, $Q = \{1, 00, 10\}$, $\pi(1) = 0.5$, $\pi(00) = 0.25$, $\pi(10) = 0.25$, $\tau(1, 0) = 0.5$, $\gamma(1, 0) = 10$, $\tau(1, 1) = 0.5$, $\gamma(1, 1) = 1$, $\tau(00, 0) = 0.75$, $\gamma(00, 0) = 00$, $\tau(00, 1) = 0.25$, $\gamma(00, 1) = 1$, $\tau(10, 0) = 0.25$, $\gamma(10, 0) = 00$, $\tau(10, 1) = 0.75$, $\gamma(10, 1) = 1$.

СУФФИКСНЫЕ ДЕРЕВЬЯ ПРЕДСКАЗАНИЯ

Определение 19 *Суффиксное дерево предсказания* (СДП) R - дерево степени $|\Sigma|$. Для каждого $\sigma \in \Sigma$ из каждой внутренней вершины выходит одно ребро, помеченное этим символом σ. Каждая вершина помечена парой (s, γ_s), где s - строка, соответствующая пути от данной вершины к корню дерева, а $\gamma_s : \Sigma \to [0, 1]$ - р.в. следующего символа: $\sum_{\sigma \in \Sigma} \gamma_s(\sigma) = 1$.

Пример 4 Пусть $\Sigma = \{0, 1\}$, $R = \{\lambda, 0, 00, 10, 1\}$, $\gamma_\lambda(0) = 0.5$, $\gamma_\lambda(1) = 0.5$, $\gamma_0(0) = 0.5$, $\gamma_0(1) = 0.5$, $\gamma_{00}(0) = 0.75$, $\gamma_{00}(1) = 0.25$, $\gamma_{10}(0) = 0.25$, $\gamma_{10}(1) = 0.75$, $\gamma_1(0) = 0.5$, $\gamma_1(1) = 0.5$.

РАСПРЕДЕЛЕНИЕ СТРОК, ПОРОЖДАЕМЫЕ СДП

Определение 20 Положим $q^0 = \lambda$ и s^j - строка, соответствуюшая вершине при проходе из корня по последовательности $\sigma_j \sigma_{j-q}$ для $1 \leq j < l$ Доказать, что вероятность того, что R породит строку $\sigma_1, \ldots, \sigma_l$ равна:

$$P_R(\sigma_1, \ldots, \sigma_l) = \prod_{i=1}^{l} \gamma_{q^{i-1}}(\sigma_i). \tag{14}$$

где $q^i = \tau(q^{i-1}, \sigma_i)$

Задача 8 *Доказать, что распределения на строках, порождаемые СВА и СДП из предыдущих примеров, совпадают.*

ПРИМЕРЫ ОБУЧЕНИЯ

М.В.ЕГОРОВА (РГГУ, 2006): Японские свечи на ежедневных ценах акций РТС

М.А.ШАРМАНОВА (РГГУ, 2007): Диаграммы „крестики-нолики“ на ежедневных ценах акций РТС

НОРМАЛЬНЫЕ СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Определение 21 $(\mu\ ,\sigma^2)$-*нормальная* случайная величина - это с.в. с плотностью $f_{\mu,\sigma^2}(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma}$.

Утверждение 5 Число μ является математическим ожиданием $\mathbf{E}(X)$, а число σ^2 - дисперсией $\mathbf{D}(X)$.

Задача 9 *Проверить, что график $y = f_{\mu,\sigma^2}(x)$ имеет максимум в точке μ, а точки перегиба в $\mu - \sigma$ и $\mu + \sigma$.*

Утверждение 6 Если X_1 и X_2 - независимые $(\mu_1\ ,\sigma_1^2)$- и $(\mu_2\ ,\sigma_2^2)$-нормальные с.в., то $X_1 + X_2$ - $(\mu_1 + \mu_2\ ,\sigma_1^2 + \sigma_2^2)$-нормальная с.в.

ЛОГНОРМАЛЬНЫЕ С.В.

Определение 22 Если X - (μ ,σ^2)-нормальная с.в., то $Z = \exp(X)$ называется (μ ,σ^2)-*логнормальной* с.в. Она принимает только положительные значения.

Задача 10 *(μ ,σ^2)-логнормальная с.в. Z имеет плотность $f^Z_{\mu,\sigma^2}(z) = \frac{e^{-(\ln(z)-\mu)^2/2\sigma^2}}{\sqrt{2\pi}\sigma z}$ для $z \in (0, \infty)$.*

Утверждение 7 (μ ,σ^2)-логнормальная с.в. Z имеет $\mathbf{E}(X) = e^{\mu+\sigma^2/2}$ и $\mathbf{D}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$.

ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА

Теорема 8 Пусть X_1**,** $\ldots$**,** X_n**,** $\ldots$ **- н.о.р. с.в. с** $\mathbf{E}(X_j) = \mu$**,** $\mathbf{D}(X_j) = \sigma^2 > 0$**. Положим** $S_n = X_1 + \ldots + X_n$**. Тогда для всех** $-\infty < a < b < \infty$

$$\lim_{n \to \infty} \mathbf{P}\left(a \leqslant \frac{S_n - n\mu}{\sqrt{n}\sigma} \leqslant b\right) = \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx. \tag{15}$$

ОПРЕДЕЛЕНИЕ ВИНЕРОВСКОГО ПРОЦЕССА

Определение 23 Случайный процесс $W = \{W(t), t \geq 0\}$ называется *винеровским процессом*, если
$W(0) = 0$ почти наверное.
$W(t_0), W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})$ независимы для любых $0 = t_0 < t_1 < \ldots < t_n$.
$W(t) - W(s)$ - $(0, t-s)$-нормальная с.в.
$W(t)$ - непрерывная функция почти наверное.

Задача 11 *Доказать, что* $cov(W(s), W(t)) = \min\{s, t\}$.

ФУНКЦИИ ХААРА

Определение 24 *Функции Хаара* $H_k(t), t \in [0,1], k = 1, 2, \ldots$ определяются так:
$H_1(t) \equiv 1$, $H_2(t) = 1_{[0,1/2]}(t) - 1_{(1/2,1]}(t)$.
$H_k(t) = 2^{n/2}(1_{I_{n,k}}(t) - 1_{J_{n,k}}(t))$ для $2^n < k \leq< 2^{n+1}$,
где $I_{n,k} = [a_{n,k}, a_{n,k} + 2^{-n-1}]$, $I_{n,k} = (a_{n,k} + 2^{-n-1}, a_{n,k} + 2^{-n}]$,
$a_{n,k} = 2^{-n}(k - 2^n - 1)$, $(n \in N)$.

Задача 12 *Доказать, что система функций Хаара - ортонормирована относительно скалярного произведения* $L^2[0,1]$*:*

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt \tag{16}$$

ПОЛНОТА СИСТЕМЫ ХААРА

Утверждение 9 Система Хаара $H_k(t), t \in [0,1], k = 1, 2, \dots$ полна в $L^2[0,1]$, т.е. для любой функции f

$$f = \sum_{k=1}^{\infty} \langle f, H_k \rangle H_k \tag{17}$$

Задача 13 *Доказать равенство Парсеваля:*

$$\langle f, g \rangle = \sum_{k=1}^{\infty} \langle f, H_k \rangle \langle H_k, g \rangle . \tag{18}$$

ЛЕММА О РЯДАХ ШАУДЕРА

Определение 25 *Функции Шаудера* $S_k(t), t \in [0,1], k = 1, 2, \dots$ - это первообразная функций Хаара $S_k(t) = \int_0^t H_k(y)dy$.

Лемма 1 Пусть числовая последовательность $\{a_k\}$ такова, что $a_k = O(k^\varepsilon)$ при $k \to \infty$ для некоторого $\varepsilon < 1/2$. Тогда ряд $\sum\limits_{k=1}^{\infty} a_k S_k(t)$ сходится равномерно на $[0,1]$ и, следовательно, задает непрерывную на $[0,1]$ функцию.

ПОСТРОЕНИЕ ВИНЕРОВСКОГО ПРОЦЕССА

Теорема 10 Пусть $\{\zeta_k\}$ **- последовательность независимых** $(0,1)$**-нормальных с.в. Положим для** $t \in [0,1]$

$$W(t) = \sum_{k=1}^{\infty} \zeta_k S_k(t). \tag{19}$$

Тогда $W = \{W(t), t \geq 0\}$ **- винеровский процесс на** $[0,1]$**.**

Теорема 11 Пусть $\{W_k\}$ **- последовательность независимых случайных функций** $[0,1]$**. Положим для** $t \in [0,\infty)$

$$W(t) = \begin{cases} W_1(t) & t \in [0,1) \\ \sum_{j=1}^{k} W_j(1) + W_{k+1}(t-k) & t \in [k,k+1) \end{cases} \tag{20}$$

Тогда W **- винеровский процесс на** $[0,\infty)$**.**

ВСПЛЕСКИ - ВЕЙВЛЕТЫ

Определение 26 *Всплеском* называют функцию $\psi : R \to C$, удовлетворяющую условиям:

$$\psi \in L^2, t \cdot \psi(t) \in L^1, \|\psi\| = 1, \int\limits_{-\infty}^{\infty} \psi(t)dt = 0 \tag{21}$$

Пример 5 $H_2(t) = 1_{[0,1/2]}(t) - 1_{(1/2,1]}(t)$ - всплеск Хаара

РАЗЛОЖЕНИЕ ПО ВСПЛЕСКАМ

Определение 27 Для фиксированного всплеска $\psi : R \to C$ коэффициенты разложения определяются следующим образом:

$$Wf(a,b) = \frac{1}{|a|^{1/2}} \int f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \tag{22}$$

Задача 14 *Доказать, что для всплеска Хаара:*

$$Wf(a,b) = \frac{|a|^{1/2}}{2} \left(\frac{2}{a} \int\limits_{b}^{b+a/2} f(t)dt - \frac{2}{a} \int\limits_{b+a/2}^{b+a} f(t)dt \right) \tag{23}$$

ФОРМУЛЫ ОБРАЩЕНИЯ

Теорема 12 Пусть x - точка непрерывности функции f. Тогда верна формула обращения:

$$f(x) = \frac{1}{C_\psi} \iint Wf(a,b) \frac{\overline{\psi\left(\frac{t-b}{a}\right)}}{|a|^{1/2}} \frac{da \cdot db}{|a|^2}, \qquad (24)$$

где константа C_ψ зависит только от всплеска ψ.

ЛИТЕРАТУРА

Булинский А.В., Ширяев А.Н. Теория случайных процессов. М.: ФИЗМАТЛИТ, 2003

Бухараев Р.Г. Основы теории вероятностных автоматов. М.: Наука, 1985

Вапник В.Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979

Добеши И. Десять лекций по вейвлетам. Ижевск: РХД, 2001

Ширяев А.Н. Основы стохастической финансовой математики. Т. 1,2. М: Фазис, 2004