

Получение статистических характеристик информативных признаков сигналов в автономных информационных системах с использованием нейросетей

03, март 2014

DOI: 10.7463/0314.0699228

профессор, д.т.н. Хохлов В. К., профессор, д.т.н. Лихоеденко К. П.

УДК 621.396.96

Россия, МГТУ им. Н.Э. Баумана

khokhlov2010@yandex.ru

wave@sm.bmstu.ru

Введение

Одной из актуальных является задача существенного повышения помехоустойчивости автономных информационных и управляющих систем (АИС) ближней локации (БЛ) в условиях интенсивных искусственных активных и пассивных помех. В ближней локации случайные сигналы на входе и помехи имеют большой динамический диапазон амплитудных, частотных и временных характеристик и ярко выраженный нестационарный характер. Информативные параметры сигналов часто являются нецентрированными случайными величинами или процессами на ограниченном интервале наблюдения, для которых априорно не известны математические ожидания, оценить которые по нестационарной реализации также не представляется возможным. В этих условиях невозможно применить традиционные методы статистических решений, требующие вычисления и обработки центрированных параметров сигналов.

Всеим отмеченным выше требованиям удовлетворяет подход к созданию принципиально новых АИС, основанный на обработке в трактах АИС нецентрированных параметров сигналов и помех, на использовании в качестве априорной информации начальных моментов случайных процессов и на применении нейросетевых технологий для исследования и проектирования АИС нового поколения [1]. В зарубежной литературе [2-4] рассматриваются вопросы применения нейросетей и линейной регрессии для распознавания объектов в радиолокационных системах, однако при этом задача решается в рамках классической теории (в том числе дискриминантного анализа) применительно к центрированным параметрам информативных признаков сигналов.

Поскольку блок принятия решения АИС БЛ в настоящее время реализуется, как правило, на основе дискретной или цифровой обработки сигналов, то в качестве первичных признаков часто используются отсчеты сигналов в дискретном времени (по Котельникову) или дискретные числовые наборы.

Общая аналитическая зависимость входного сигнала без постоянной составляющей может быть представлена в виде

$$x(t) = E(t) \cos \left[\int_0^t \omega(z) dz + \varphi(t) \right],$$

где $E(t)$ – огибающая; $\omega(t)$ – мгновенная частота; $\varphi(t)$ – случайная фаза.

Если информация о физических и механических признаках объектов заключена в статистических характеристиках огибающей, случайной фазы или мгновенной частоты и нестационарность входных сигналов в БЛ вызвана изменением во времени одного или нескольких из этих информативных признаков, то, как показано в [1, стр.43-47,96-99], в АИС БЛ целесообразно не проводить обработку сигнала в целом, а выделять перечисленные информативные признаки. Эти признаки в общем случае представляют собой случайные нестационарные процессы, для которых априорно не известны математические ожидания, оценка последних часто не представляется возможной из-за высокого быстродействия систем. Отметим, что в АИС информация о мгновенной частоте и случайной фазе часто получается в результате обработки интервалов между нулями входных реализаций [1, стр.97].

Важными вопросами при разработке АИС БЛ являются вопросы выбора информативных признаков и вопросы сокращения размерности векторов признаков.

Цель выбора информативных признаков сигналов состоит в выборе таких признаков, которые являются наиболее эффективными с точки зрения разделимости классов.

Научная новизна

В теории распознавания образов для выбора признаков и сокращения размерности векторов признаков широко используются разложения Карунена–Лоэва, его дискретный аналог – метод главных компонент и метод множественного дискриминантного анализа. Эти методы оперируют центральными (ковариационными) моментами случайных отсчетов сигналов, последние два метода предполагают нахождение собственных векторов и собственных чисел по заданной ковариационной матрице.

В многоканальных системах БЛ часто информация о пространственно – геометрических и физических признаках объектов заключена в соотношениях между детерминированными составляющими параметров сигналов при каждом из конкретных условий встреч и задачи по построению АИС часто приходится решать в условиях неизвестных математических ожиданий, поэтому применение перечисленных выше классических методов часто оказывается затруднительным.

В [5] исследован метод выбора и сокращения размерности вектора входной реализации, основанный на начальных регрессионных представлениях. Для реализации метода необходимо вычисление остаточных средних квадратических значений множественных начальных регрессионных представлений и коэффициентов начальной регрессии.

Научная новизна определяется применением нейросетевых алгоритмов для эффективного решения задач выбора информативных признаков и определения параметров регрессионных алгоритмов в условиях вырожденности или плохой обусловленности данных при неизвестных математических ожиданиях и ковариационных матрицах.

Постановка задачи

В [1, стр. 161] показано что, что при обнаружении сигналов на фоне белого шума в регрессионных алгоритмах необходимо вычислять квадраты ошибок множественных начальных представлений вида

$$\sum_{i=1}^n \left(\bar{x}_i - \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik} \bar{x}_k \right)^2,$$

где: β_{ik} – коэффициенты начальной регрессии; $\bar{x}_k$ – оценки нецентрированных параметров сигналов. Нецентрированная регрессионная оценка случайного параметра $\bar{x}_i$ есть

$$\hat{\bar{x}}_i = \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik} \bar{x}_k. \quad (1)$$

В [1, стр.159] обоснованы регрессионные обнаружения и распознавания сигналов на фоне полосового белого шума

$$\frac{1}{2f_B S_0} \sum_{i=1}^n \bar{x}_i^2 - \frac{1}{4} \sum_{i=1}^n \frac{1}{\bar{\Psi}_{0ci}^2} \left(\bar{x}_i - \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik} \bar{x}_k \right)^2 \geq \gamma_1^*, \quad (2)$$

$$-\sum_{i=1}^n \frac{1}{\bar{\Psi}_{0ci}^2} \left(\bar{x}_i - \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik}^c \bar{x}_k \right)^2 + \sum_{i=1}^n \frac{1}{\bar{\Psi}_{0ni}^2} \left(\bar{x}_i - \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik}^n \bar{x}_k \right)^2 \geq 0, \quad (3)$$

где: f_B - верхняя частота в полосе пропускания тракта; S_0 - спектральная плотность белого шума; $\bar{x}_i$ и $\bar{x}_k$ - оценки информативных параметров; $\bar{\Psi}_{0ci}^2$ и $\bar{\Psi}_{0ni}^2$ - остаточные средние оценок случайных параметров соответственно для сигнала и помехи.

На основании [1, стр.159] (2) и (3) записаны через остаточные средние значения квадратов $\bar{\Psi}_{0i}^2$ множественных начальных регрессионных представлений. Как видно из (1), (2) и (3) в алгоритмах вычисляются отношения квадрата ошибки множественных регрессионных представлений координат случайного входного вектора к соответствующим остаточным средним значениям квадратов $\bar{\Psi}_{0i}^2$.

Рассмотрим однослойную нейронную сеть (рисунок 1), в которой каждый элемент из множества входов X отдельным весом соединен с каждым искусственным нейроном, а каждый нейрон выдает взвешенную сумму в сеть. Матрица весов W имеет m строк и n столбцов, где m - число входов, а n - число нейронов. Вычисление выходного вектора Y , компонентами которого являются выходы нейронов, сводится к матричному умножению

$$Y = XW. \quad (4)$$

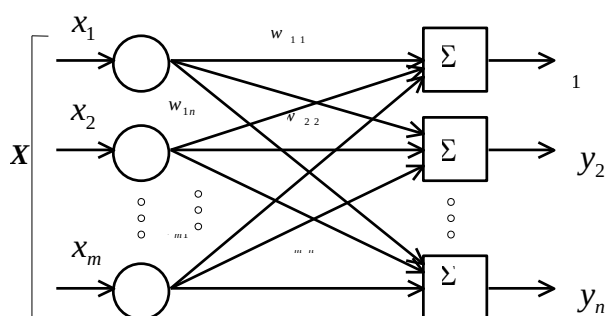


Рис. 1. Структурная схема однослойной нейронной сети без нулевых весов

Как видно из рисунка 1 при использовании в качестве входов (признаков)

нецентрированных параметров сигналов в нейросети осуществляется обработка нецентрированных значений сигналов.

Исключим из рассматриваемой сети прямые связи, т.е. положим $w_{ii} = 0$
 $i=1,2,\dots,m$ и рассмотрим общий случай сети с m нейронами.

Потребуем на каждом i -ом выходе нейрона восстановления i -го входного сигнала с минимальной ошибкой, т.е. в рассматриваемом случае при обработке нецентрированных параметров сигналов, с минимумом среднего значения квадрата $\overline{\Psi}_{yi}^2$. Произведем невырожденное преобразование вектора X

$$Y = XW,$$

т.е. найдем матрицу W так, чтобы остаточная сумма квадратов

$$\Psi\Psi^T = \Psi^T\Psi = M[(Y - WX)(Y - WX)^T] \quad (5)$$

была минимальной.

Дифференцируя остаточную сумму квадратов $\Psi\Psi^T$ по W и приравнявая её нулю, для положительно определённой и симметричной матрицы начальных корреляционных моментов K (аналогично [1, стр.40-41]) получим

$$w_{ik} = -\Lambda_{ik} / \Lambda_{ii}, \quad (6)$$

где $\Lambda = K^{-1}$, K – матрица корреляционных (начальных) моментов.

Как следует из (6) весовые коэффициенты w_{ik} в рассматриваемом случае представляют собой коэффициенты начальной регрессии $w_{ik} = \beta_{ik}$ в уравнении начальной регрессии (1), [1, стр.42], минимизирующем среднее значение квадрата ошибки $\overline{\Psi}_{yi}^2$ множественного начального регрессионного представления на каждом выходе нейронной сети

$$M\left[\left(\bar{x}_i - \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik} \bar{x}_k\right)^2\right] = \overline{\Psi}_{yi}^2,$$

где: $M[\]$ – оператор математического ожидания; $\overline{\Psi}_{yi}^2$ – среднее значение квадрата оценки

i -го параметра; $\beta_{ik} = -\frac{\Lambda_{ik}}{\Lambda_{ii}}$ - начальные коэффициенты регрессии; $\Lambda_{ik}, \Lambda_{ii}$ — элементы матрицы Λ , обратной матрице корреляционных моментов K , $\Lambda = K^{-1}$.

Нейросеть осуществляет линейную обработку по алгоритму

$$y_i = \sum_{k=1, k \neq i}^n w_{ik} \bar{x}_k = \sum_{\substack{k=1 \\ k \neq i}}^n \beta_{ik} \bar{x}_k,$$

где $\bar{x}_k$ - элемент входного вектора.

Обоснование способа получения априорной информации в условиях плохой обусловленности данных

Проблема обусловленности данных возникает при анализе разделимости классов и сокращении размерности векторов информативных признаков в АИС БЛ. Плохая обусловленность данных приводит к значительному накоплению погрешности при численном решении задач.

Обусловленность является одним из важнейших свойств любой вычислительной задачи наряду с корректностью задачи (существованием, устойчивостью и единственностью ее решения).

Теоретически наличие у задачи устойчивости означает, что ее решение может быть найдено со сколь угодно малой погрешностью, если только гарантировать, что погрешности входных данных достаточно малы. Однако на практике погрешности входных данных не могут быть сколь угодно малыми, точность их ограничена. Обусловленность вычислительной задачи [6] есть чувствительность ее решения к малым погрешностям входных данных. Задачу называют хорошо обусловленной, если малым погрешностям входных данных отвечают малые погрешности решения, и плохо обусловленной, если возможны сильные изменения решения. Часто оказывается возможным ввести количественную меру степени обусловленности - число обусловленности. Эту величину можно интерпретировать как коэффициент возможного возрастания погрешностей в решении по отношению к вызвавшим их погрешностям входных данных.

Рассмотрим исходную систему линейных алгебраических уравнений

$$Ax = f \tag{7}$$

с квадратной матрицей A порядка n . Для каждого n -мерного вектора f решение x задачи существует тогда и только тогда, когда $\det A \neq 0$. В этом случае можно определить матрицу A^{-1} ,

обратную матрицу $\mathbf{A}$, и записать решение в виде:

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{f}.$$

Рассмотрим “возмущенную систему”

$$\mathbf{A} \tilde{\mathbf{x}} = \tilde{\mathbf{f}}, \quad (8)$$

которая отличается от (7) правой частью. Будем предполагать, что в матрицу $\mathbf{A}$ возмущений не вносится. Нас интересует, насколько сильно может измениться решение $\mathbf{x}$ в результате изменения правой части. Обозначим

$$\delta \mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}, \quad \delta \mathbf{f} = \tilde{\mathbf{f}} - \mathbf{f}.$$

Говорят, что система (8) устойчива по правой части, если при любых $\tilde{\mathbf{f}}$ справедлива оценка

$$\|\delta \mathbf{x}\| \leq M_1 \|\delta \mathbf{f}\|, \quad (9)$$

где $M_1 > 0$ - постоянная, не зависящая от правых частей $\tilde{\mathbf{f}}$, $\|\mathbf{A}\|$ – норма матрицы.

Легко показать, что если $\det \mathbf{A} \neq 0$, то система (8) устойчива по правой части. Действительно, уравнение для погрешности имеет вид

$$\mathbf{A}(\delta \mathbf{x}) = \delta \mathbf{f},$$

откуда получаем

$$\begin{aligned} \delta \mathbf{x} &= \mathbf{A}^{-1}(\delta \mathbf{f}), \\ \|\delta \mathbf{x}\| &\leq \|\mathbf{A}^{-1}\| \|\delta \mathbf{f}\|, \end{aligned}$$

т.е. выполняется неравенство (9) с константой $M_1 = \|\mathbf{A}^{-1}\|$. Заметим, что чем ближе к нулю определитель матрицы $\mathbf{A}$, тем больше M_1 и, следовательно, тем сильнее погрешность правой части может исказить искомое решение.

Рассмотрим число обусловленности матрицы, так как именно матрица корреляционных моментов представляет для нас непосредственный интерес. Применяя общее определение относительного числа обусловленности, можно сказать, что число обусловленности матрицы есть мера возрастания относительной погрешности решения задачи, в которой фигурирует данная матрица. Следовательно, обусловленность матрицы корреляционных моментов в нашем случае непосредственно влияет на точность вычисления коэффициентов начальной множественной регрессии и остаточного среднего значения квадрата.

В оценку (9) входят абсолютные погрешности решения $\delta \mathbf{x}$ и правой части $\delta \mathbf{f}$. При решении системы (8) на ЭВМ с плавающей запятой более естественными характеристиками являются относительные погрешности

$$\frac{\|\delta \mathbf{f}\|}{\|\mathbf{f}\|}, \quad \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}.$$

Получим оценку, выражающую относительную погрешность решения через относительную погрешность правой части. Для этого используем неравенство

$$\|\mathbf{f}\| \leq \|\mathbf{A}\| \|\mathbf{x}\|.$$

В результате мы приходим к требуемой оценке:

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq M_A \frac{\|\delta \mathbf{f}\|}{\|\mathbf{f}\|},$$

где

$$M_A = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|.$$

Число M_A , входящее в эту оценку, называется числом обусловленности матрицы $\mathbf{A}$ и характеризует степень зависимости относительной погрешности решения от относительной погрешности правой части. Матрицы с большим числом обусловленности M_A называются плохо обусловленными матрицами. При численном решении систем с плохо обусловленными матрицами возможно сильное накопление погрешностей.

Подчеркнем, что число обусловленности не связано с каким-либо численным алгоритмом, а характеризует только свойство системы (8).

Число M_A чаще обозначают как $\text{cond}(\mathbf{A})$, в дальнейшем будем придерживаться последнего обозначения.

Проблема обусловленности матрицы корреляционных моментов информативных признаков возникает при решении задачи нахождения множественных коэффициентов начальной регрессии (МКНР) и остаточных средних значений квадратов множественных регрессионных представлений. Для получения МКНР и нахождения остаточных средних значений квадратов требуется обращать матрицу корреляционных моментов случайного входного вектора. Обусловленность матриц рассмотрим на примере обусловленности матрицы корреляционных моментов интервалов между нулями случайного процесса. Выбор этой матрицы основан на возможности получения ее

теоретически. Использование вычисленного теоретически значения матрицы позволит получить более достоверные результаты.

Теоретически элементы матрицы **K** ковариационных моментов интервалов между нулями стационарного случайного процесса можно получить по формулам [1, стр. 50]:

$$K_{ik} = C(\tau_i, \tau_k) + \mu_{\tau_i} \mu_{\tau_k},$$

$$C(\tau_i, \tau_k) = \frac{1}{\omega_0^2} \left\{ 2B_\varphi \left[(l-k) \frac{\pi}{\omega_0} \right] - B_\varphi \left[(l-k+1) \frac{\pi}{\omega_0} \right] - B_\varphi \left[(l-k-1) \frac{\pi}{\omega_0} \right] \right\},$$

$$\mu_{\tau_1} = \mu_{\tau_k} = \frac{\pi}{\omega_0}.$$

где: ω_0 – средняя частота энергетического спектра сигнала; B_φ – корреляционная функция фазы процесса. Корреляционную функцию фазы B_φ необходимо определить при фиксированных значениях аргумента $\tau = K \frac{\pi}{\omega_0}$, где $\mu_\tau = \frac{\pi}{\omega_0}$ – математическое ожидание случайной длительности интервалов между нулями сигнала, K – сдвиг между сечениями (рис.2).

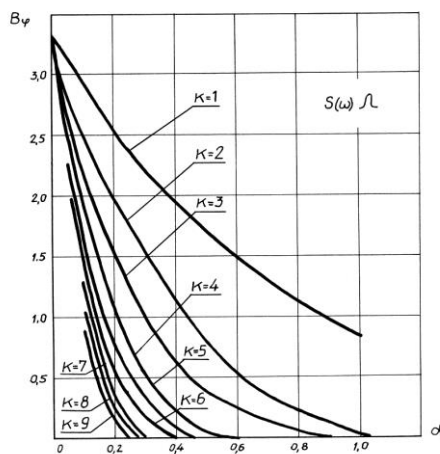


Рис. 2. Зависимость корреляционной функции фазы стационарного случайного процесса от относительной ширины полосы гауссова энергетического спектра при различных сдвигах K между сечениями

Коэффициенты начальной регрессии интервалов между нулями могут быть определены как

$$\beta_{\tau_1 / \tau_k} = \beta_{\tau_k / \tau_1} = \frac{C(\tau_1, \tau_k) + \mu_{\tau_1} \mu_{\tau_2}}{\sigma_{\tau}^2 + \mu_{\tau}^2} =$$

$$\left\{ 2B_{\varphi} \left[(k-1) \frac{\pi}{\omega_0} \right] - B_{\varphi} \left(k \frac{\pi}{\omega_0} \right) - B_{\varphi} \left[(k-2) \frac{\pi}{\omega_0} \right] + \pi^2 \right\}$$

$$\left\{ 2 \left[B_{\varphi}(0) - B_{\varphi} \left(\frac{\pi}{\omega_0} \right) \right] + \pi^2 \right\}^{-1},$$

где: математическое ожидание случайных длительностей интервалов между нулями и их дисперсия определяются выражениями:

$$\mu_{\tau_1} = \mu_{\tau_k} = \frac{\pi}{\omega_0}; \quad \sigma_{\tau}^2 = \frac{2}{\omega_0^2} \left[B_{\varphi}(0) - B_{\varphi} \left(\frac{\pi}{\omega_0} \right) \right].$$

Как показано в [1, стр.52] ковариационные моменты интервалов между нулями зависят от относительной ширины полосы спектра процесса $\alpha = \Delta\omega / \omega_0$, где $\Delta\omega$ - эффективная ширина полосы энергетического спектра случайного процесса. Это дает возможность исследовать ее обусловленность не только в зависимости от размерности входного вектора, но и в зависимости от относительной ширины полосы спектра сигнала.

Вычислительный эксперимент

Вычислительный эксперимент проводился на ПК типа Intel L(R) Core™ Duo CPU E7300@2,66ГГц, 3,25 ОЗУ.

На рисунках 3, 4, 5 приведены графики зависимостей числа обусловленности матрицы **K** корреляционных моментов интервалов между нулями стационарного случайного процесса в зависимости от относительной ширины полосы α при различных $n=const$, где n - размерность входного вектора.

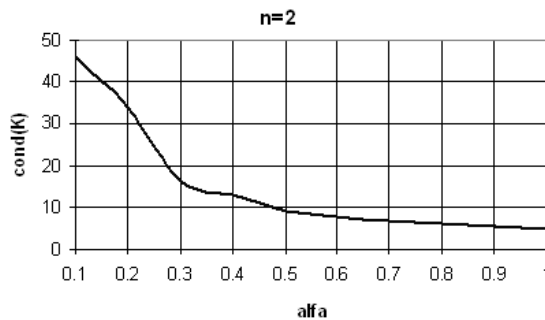


Рис. 3. Зависимость числа обусловленности матрицы **K** от α при $n=2$ для сигнала с прямоугольным спектром

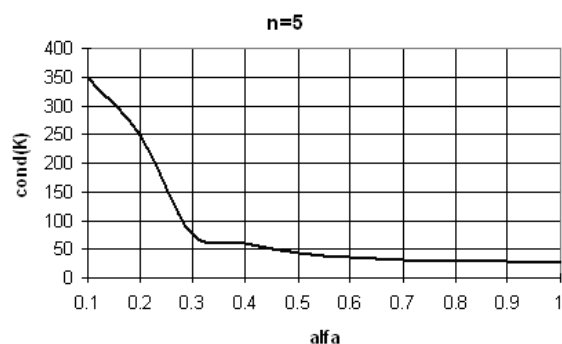


Рис. 4. Зависимость числа обусловленности матрицы $\mathbf{K}$ от α при $n=5$ для сигналов с прямоугольным спектром

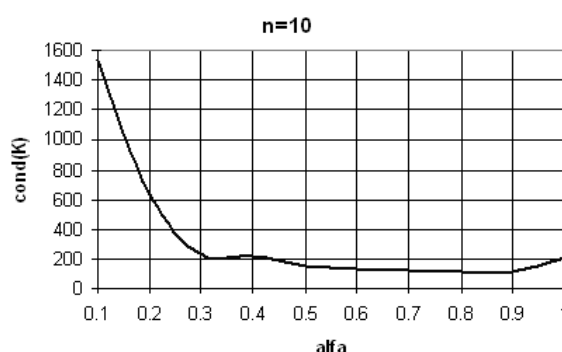


Рис. 5. Зависимость числа обусловленности матрицы $\mathbf{K}$ от α при $n=10$ для сигнала с прямоугольным спектром

Из приведенных зависимостей видно, что при уменьшении относительной ширины полосы энергетического спектра стационарного случайного процесса число обусловленности матрицы корреляционных моментов интервалов между нулями резко возрастает. При моногармоническом процессе матрица $\mathbf{K}$ вырождена, следовательно, она не имеет обратной и вычислить коэффициенты начальной регрессии через элементы обратной матрицы не представляется возможным. Опираясь на определение числа обусловленности матрицы, можно сделать вывод, что погрешности вычислений с помощью таких матриц возрастают на несколько порядков, что является недопустимым.

На рисунках 6 – 8 приведены графики зависимостей числа обусловленности матрицы корреляционных моментов интервалов между нулями стационарного случайного процесса в зависимости от размерности входного вектора n при различных α .

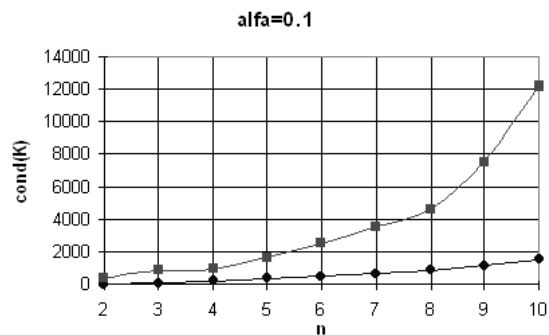


Рис. 6. Зависимость числа обусловленности матрицы $\mathbf{K}$ от n при $\alpha=0.1$ для сигналов с прямоугольным (o) и гауссовым ($\square$) спектром

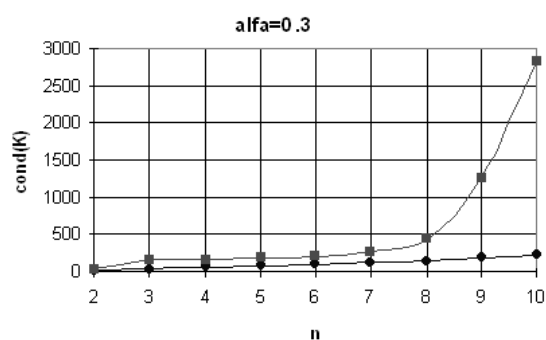


Рис. 7. Зависимость числа обусловленности матрицы $\mathbf{K}$ от n при $\alpha=0.3$ для сигналов с прямоугольным (o) и гауссовым ($\square$) спектром

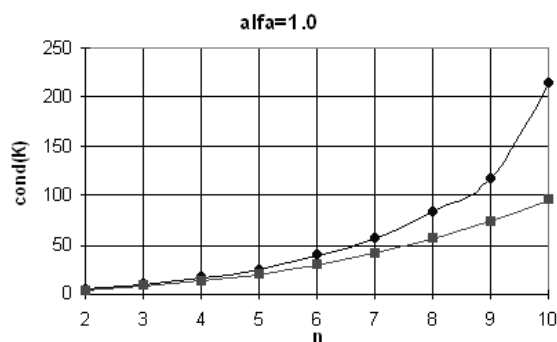


Рис. 8. Зависимость числа обусловленности матрицы $\mathbf{K}$ от n при $\alpha=1.0$ для сигналов с прямоугольным (o) и гауссовым ($\square$) спектром

Из графиков видно, что с увеличением размерности входного вектора число обусловленности резко возрастает. Этот рост тем больше, чем меньше относительная полоса спектра сигнала α .

Таким образом, в случаях, когда размерность входного вектора превышает 10 отсчетов и $\alpha < 0.3$ вычисление коэффициентов множественной регрессии через элементы обратной матрицы может повлечь за собой сильное накопление ошибок (более чем на три порядка). Полученные таким образом результаты с трудом можно считать достоверными.

В качестве альтернативного подхода предложено получить коэффициенты регрессии при помощи нейронной сети.

Рассматривалась однослойная нейронная сеть с n входами (рисунок 1), обрабатывающая нецентрированные значения интервалов между нулями нормального стационарного случайного процесса.

Сеть обучалась на математической модели сигнала по алгоритму обратного распространения [7].

Сглаженная зависимость величины ошибки на выходе сети от номера итерации приведена на рисунке 9.

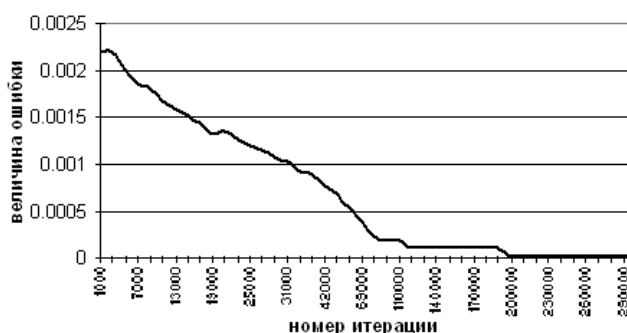


Рис. 9. Зависимость величины ошибки при обучении сети от номера итерации

Как показывают результаты расчетов (рисунок 8), величина ошибки достигает приемлемо малой величины ($\varepsilon \approx 10^{-5}$) уже через 100000 итераций.

Полученные в качестве весов такой сети коэффициенты множественные коэффициенты регрессии первого интервала на остальные при общей размерности задачи равной семи для различных α приведены на рисунках 10– 11.



Рис. 10. Коэффициенты множественной регрессии 1-го интервала на 6 остальных при $\alpha = 0.1$

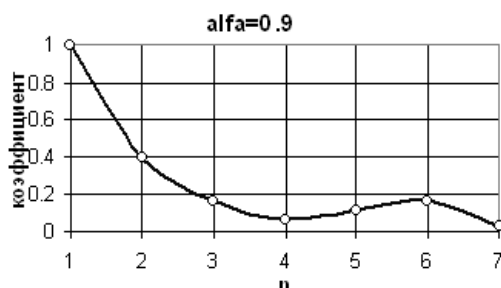


Рис. 11. Коэффициенты множественной регрессии 1-го интервала на 6 остальных при $\alpha = 0.9$

Применение предлагаемого нейросетевого подхода позволяет получить оценки остаточных средних значений квадратов регрессионных представлений $\overline{\Psi}_{xi}^2$ и коэффициентов множественной начальной регрессии интервалов между нулями сигналов и помех как весовые коэффициенты W_{ij} , получаемые при обучении нейросети с использованием метода обратного распространения. Результаты исследований показали, что по предлагаемой методике легко могут быть вычислены требуемые статистические характеристики даже при нормированных коэффициентах корреляции r_{ik} параметров сигналов стремящихся к единице.

В [1, стр. 247] показано, что при $r_{ik} = 1$, $\beta_{ik} = \frac{1}{n-1}$, где n - размерность входного вектора. Это позволило осуществить проверку правильности выполненных расчетов. График, приведенный на рисунке 10 подтверждает тот факт, что при $\alpha \rightarrow 0$ все КМНР стремятся к одинаковому значению: $\beta_{ik} = \frac{1}{n-1}$, где n - размерность входного вектора. В нашем случае при $n=7$ $\beta_{1i} = \frac{1}{6} = 0.16$. С увеличением относительной ширины полосы усиливается вклад первых двух отсчетов.

Важно отметить, что нейронная сеть является однослойной с линейной функцией активации, что значительно сокращает затраты времени на ее обучение и требования к ресурсам

вычислительной машины. После обучения нейронной сети ее веса представляют собой коэффициенты множественной регрессии (в рассматриваемом случае начальные).

Заключение

В статье рассмотрены вопросы получения статистических характеристик векторов входных реализаций сигналов и помех для решения задач выбора информативных признаков, сокращения размерностей векторов признаков и обоснования алгоритмов обнаружения и распознавания сигналов от помех в условиях неизвестных математических ожиданий. В основу положен метод начальных регрессионных представлений. Показано, что при плохой обусловленности матриц корреляционных моментов информативных признаков применение предлагаемого нейросетевого подхода позволяет получить оценки коэффициентов множественной начальной регрессии параметров признаков и остаточных средних значений квадратов регрессионных представлений $\overline{\Psi}_{xi}^2$ как весовые коэффициенты W_{ij} , и их дисперсии, получаемые при обучении нейросети. Результаты исследований показали, что по предлагаемой методике легко могут быть вычислены требуемые статистические характеристики даже при плохо обусловленных матрицах ковариационных моментов, когда нормированные коэффициенты корреляции параметров признаков сигналов стремятся к единице.

Список литературы

1. Автономные информационные и управляющие системы: в 4 т. Т. 1 / Ю.М. Астапов, А.Б. Борзов, А.К. Ефремов [и др.]; под ред. А.Б. Борзова. М.: ООО НИЦ «Инженер», ООО «Онико-М», 2011. С. 5-293.
2. Lin Yun L., Jing-chao L. Radar signal recognition algorithms based on neural network and grey relation theory // Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). Vol. 2. 2011. P. 1482-1485. DOI: 10.1109/CSQRWC.2011.6037247
3. Javed A., Ejaz A., Liaqat S., Ashraf A., Ihsan M.B. Automatic target classifier for a Ground Surveillance Radar using linear discriminant analysis and Logistic regression // 2012 9th European Radar Conference (EuRAD). 2012. P. 302-305.
4. Ali S.F., Jaffar J., Malik A.S. Proposed framework of Intelligent Video Automatic Target Recognition System (IVATRs) // 2011 National Postgraduate Conference (NPC). 2011. P. 1-5. DOI: 10.1109/NatPC.2011.6136315
5. Хохлов В.К., Гулин Ю.Ю. Выбор информативных признаков в автономных информационных системах с нейросетевыми трактами обработки сигналов // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2003. № 3. С. 70-83.
6. Демидович Б.П., Марон И.А. Основы вычислительной математики. М.: Наука, 1970.

7. Оссовский С. Нейронные сети для обработки информации / Пер. с польск. И.Д. Рудинского. М.: Финансы и статистика, 2002. 344 с.

The obtaining of statistical characteristics of informative features of signals in the Autonomous information systems using neural networks

03, March 2014

DOI: **10.7463/0314.0699228**

V.K. Hohlov, K.P. Likhoedenko

Bauman Moscow State Technical University, 105005, Moscow, Russian Federation

khokhlov2010@yandex.ru

wave@sm.bmstu.ru

The article studies a neural network approach to obtain the statistical characteristics of the input vector implementations of signal and noise at ill-conditioned matrices of correlation moments to solve the problems to select and reduce the vector dimensions of informative features at detection and recognition of signals and noise based on regression methods.

A scientific novelty is determined by applying neural network algorithms for the efficient solution of problems to select the informative features and determine the parameters of regression algorithms in terms of the degeneracy or ill-conditioned data with unknown expectation and covariance matrices.

The article proposes to use a single-layer neural network with no zero weights and activation functions to calculate the initial regression characteristics and the mean-square value error of multiple initial regression representations, which are necessary to justify the selection of informative features, reduce a dimension of sign vectors and implement the regression algorithms. It is shown that when excluding direct links between the inputs and their corresponding neurons, in the training network the weight coefficients of neuron inputs are the coefficients of initial multiple regression, the error mean-square value of multiple initial regression representations is calculated at the outputs of neurons. The article considers conditionality of the problem to calculate the matrix that is inverse one for matrix of correlation moments. It defines a condition number, which characterizes the relative error of stated task.

The problem concerning the matrix condition of the correlation moment of informative signal features and noise arises when solving the problem to find the multiple coefficients of initial regression (MCIR) and the residual mean-square values of the multiple regression representations. For obtaining the MCIR and finding the residual mean-square values the matrix of correlation moments of the random input

vector should be inverted. Conditionality of matrices is considered by the example of the correlation moment matrix of the intervals between the zeros of a stationary random process. The choice of this matrix is defined by a possibility to obtain it theoretically and from MCIR calculations.

The article presents dependences representing matrix elements of correlation moments of intervals between the zeros of a stationary random process, which are off-center random values and are used in regression and neural network systems to distinguish signals from noise. It is shown that the elements of the correlation moments of the intervals between the zeros of a stationary random process depend on the correlation function of the process phase and are determined by the relative bandwidth of the process power distribution. To conduct the computational experiment a computer was used. It shows the condition number of the correlation moment matrix of the intervals between the zeros versus the relative width of the processes at different dimension of input vector for signals with rectangular and Gaussian spectra. It follows from the presented dependences that with decreasing the relative bandwidth of the power distribution of a stationary random process, the condition number of the correlation moment matrix of the intervals between the zeros dramatically increases. In cases when the dimension of the input vector exceeds 10 counts and the relative bandwidth of the power distribution is less than 0.3 the calculated coefficients of multiple regression through the elements of the inverse matrix entails the accumulation of errors by more than three orders of magnitude. The article concerns the issue to obtain the coefficients of the initial regression using neural network. It is shown that the application of the proposed neural network approach enables us to estimate the residual mean-square values of the multiple regression representations and coefficients of multiple initial regression of intervals between the signal zeros and noise as the weight factors of neuron inputs obtained when training the neural network by the back propagation method for less than 100,000 iterations. By the offered technique initial regression coefficients are calculated even at the rated correlation coefficients of the signal parameters aspiring to unit, i.e. with a degenerate matrix of initial correlation moments. The obtained results have been validated theoretically. The results presented in article can be applied with justified informative features, reduced dimension of sign vectors, and implemented regression algorithms in systems of close location the specifics of which often does not allow the centering signal parameters.

Publications with keywords: [correlation matrix](#), [informative features](#), [ill-conditioning](#), [training of the neural network](#), [the weights](#), [the regression coefficients](#)

Publications with words: [correlation matrix](#), [informative features](#), [ill-conditioning](#), [training of the neural network](#), [the weights](#), [the regression coefficients](#)

References

1. Astapov Yu.M., Borzov A.B., Efremov A.K., et al. *Avtonomnye informatsionnye i upravliaiushchie sistemy: v 4 t. T. 1* [Autonomous information and control systems: in 4 vols. Vol. 1]. Moscow, OOO NITs "Inzhener" Publ., OOO "Oniko-M" Publ., 2011, pp. 5-293. (in Russian).
2. Lin Yun L., Jing-chao L. Radar signal recognition algorithms based on neural network and grey relation theory. *Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC, vol. 2, 2011, pp. 1482-1485. DOI: 10.1109/CSQRWC.2011.6037247*
3. Javed A., Ejaz A., Liaqat S., Ashraf A., Ihsan M.B. Automatic target classifier for a Ground Surveillance Radar using linear discriminant analysis and Logistic regression. *2012 9th European Radar Conference (EuRAD)*, 2012, pp. 302-305.
4. Ali S.F., Jaffar J., Malik A.S. Proposed framework of Intelligent Video Automatic Target Recognition System (IVATRs). *2011 National Postgraduate Conference (NPC)*, 2011, pp. 1-5. DOI: 10.1109/NatPC.2011.6136315
5. Khokhlov V.K., Gulin Yu.Yu. [Informative signs selection in autonomous information systems with neuron net paths of signal processing]. *Vestnik MGTU. Ser. Priborostroenie - Herald of the Bauman MSTU. Ser. Instrument Engineering*, 2003, no. 3, pp. 70-83. (in Russian).
6. Demidovich B.P., Maron I.A. *Osnovy vychislitel'noy matematiki* [Foundations of computational mathematics]. Moscow, Nauka, 1970. (in Russian).
7. Ossovskiy S. *Neyronnye seti dlya obrabotki informatsii* [Neural networks for information processing]. Transl. from Polish. Moscow, Finansy i statistika, 2002. 344 p.