ЛИТЕРАТУРА

1. P. E. Kalman. New approach to linear filtering and prediction problems.— "Trans. ASME Basic Eng.", March 1960.
2. Ю. А. Кочетков, В. К. Томшин. Оптимальная оценка искомой величины в случае негауссовых ошибок измерения.— «Техн. кибернетика», 1968, № 5.
3. М. Аоки. Оптимизация стохастических систем. М., «Наука», 1971.
4. А. М. Якубович. Экстраполяция случайного процесса, измеряемого несколькими приборами с учетом надежности их работы.— «Техн. кибернетика», 1971, № 4.
5. В. С. Пугачев. Теория случайных процессов. М., «Энергия», 1963.
6. Р. Ли. Оптимальные оценки, определение характеристик и управления. М., «Наука», 1966.

Поступила в редакцию
4 января 1974 г.;
окончательный вариант —
3 июля 1975 г.

УДК 62-50

В. И. МИЗЮКОВ, К. К. ПАЩЕНКО,
В. В. СОКОЛОВСКИЙ, В. А. СТАНОВОВА

(Караганда)

РЕШЕНИЕ НЕКОТОРЫХ ЗАДАЧ ОПТИМИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ ЭЛЕМЕНТОВ ТЕНЗОРНОГО АНАЛИЗА

Значительным шагом на пути создания общей теории систем является, как известно, метод пространства состояний, позволяющий с единых позиций рассматривать линейные и нелинейные, стационарные и нестационарные системы, хотя решение нестационарных и нелинейных задач остается по-прежнему трудным [1, 2]. Основной математический аппарат метода пространства состояний — векторно-матричная алгебра в линейных конечномерных пространствах с введением операции «векторного дифференцирования» [2]. Несмотря на то что понятие «состояния» системы, действительно, универсально и плодотворно для всех систем, математический аппарат метода пространства состояний адекватен в основном описанию (в полной и в меньшей мере анализу) линейных стационарных многомерных систем. При математическом моделировании и анализе нестационарных систем, описании и анализе функций качества и в ряде других прикладных задач традиционный аппарат метода пространства состояний уже для описания требует таких искусственных конструкций, как блочные матрицы, составные векторы, обобщенный обратный матричный оператор и т. п., не наглядных и не удобных при исследовании систем с помощью ЭВМ.

Эти затруднения иллюстрируются на примере идентификации линейной стационарной системы, описанной в соответствии с [3] системой линейных дифференциальных уравнений:

$$\dot{x}=Ax+Bv;\ y=Hx, \tag{1}$$

где A, B, H — постоянные матрицы, y — вектор измерений, v — вектор управлений, x — вектор состояний.

Требуется идентифицировать матрицу A по результатам измерений y. Разработанные векторные итеративные градиентные процедуры идентификации имеют вид

$$c[n]=c[n-1]-\gamma[n]\nabla_c J[n-1], \tag{2}$$

где c — элементы матрицы A, сформированные в вектор; $\nabla_c I$ — градиент критерия качества идентификации, вычисленный с применением операции «векторного дифференцирования» [3].

Эффективность процедуры (2) зависит от вида критерия качества I. Если I является квадратичным функционалом в пространстве определяемых параметров, то при его минимизации становится возможным получить аналитическое выражение градиента. Если критерий качества I рассматривать в пространстве матриц A, то для элементов матрицы A получим итеративную процедуру типа (2). При этом возникает необходимость дифференцирования критерия качества I по матрице A (для линейных стационарных систем) и в общем случае по тензору. Необходимость дифференцирования по тензору возникает уже при анализе линейных нестационарных систем. Рекуррентная схема отыскания матрицы A имеет вид

$$A[n+1]=A[n]-\gamma[n]\nabla_A I[n], \tag{3}$$

где $\nabla_A I$ — градиент критерия качества I, вычисленный с применением операции матричного дифференцирования [4]. Трудности, связанные с ограниченными возможностями векторно-матричной алгебры, особенно проявляются при анализе сложных систем. Для их преодоления необходимо создать математический аппарат, адекватный задачам анализа систем (в том числе нестационарных и нелинейных), что возможно при расширении круга линейных алгебраических объектов, используемых при описании систем, т. е. привлечении к описанию систем более общих алгебраических объектов — тензоров валентности выше второй.

В общем случае задача анализа систем может быть сформулирована как исследование на экстремум критерия качества системы, представляющего собой в подавляющем большинстве случаев функционал системы, т. е. получение и решение уравнения

$$\frac{dI(A)}{dA}=0. \tag{4}$$

До настоящего времени известны лишь запись уравнения (4) и его решение в случаях, когда A — скаляр, вектор или матрица [3, 5, 6].

Необходимо отметить, что критерий качества системы I в общем случае является многочленным комитантом [7], т. е. своеобразной функцией, ставящей в соответствие некоторой последовательности заданных тензоров $a, \ldots, c$ определенных типов некоторый новый тензор $I_{i_1,\ldots,i_p}(a, \ldots, c)$ в соответствии с операциями алгебры тензоров (сложением, умножением, свертыванием, перестановкой индексов). Условимся применять для тензоров только ковариантные индексы; этого всегда можно достичь, предположив, что тензоры будут рассматриваться в евклидовом или аффинном пространствах. Очевидно, координаты многочленных комитантов выражаются в виде многочленов от координат заданных тензоров, причем коэффициенты этих многочленов не зависят от выбора координатного базиса в основном пространстве.

Получение и решение уравнения (4) требует введения операции дифференцирования функционала I по тензорной переменной A. Так как в большинстве случаев I является сложной функцией от тензорных переменных, то необходимо ввести операцию дифференцирования комитанта по тензору.

Операция дифференцирования комитантов по заданным тензорам в общем случае достаточно трудоемка. Однако для многочленных комитантов можно получить в явном виде довольно простые выражения для частных производных по тензорам переменных, если воспользоваться представлением инвариантов и целых рациональных ковариантов в виде линейной комбинации произведений заданных тензоров со свертками [7]. Тогда дифференцирование комитанта по заданному тензору сводится к дифференцированию отдельных произведений заданных тензоров со свертками.

Для дифференцирования по тензору произведения заданных тензоров со свертками введем следующее правило нахождения частных производных:

$$\frac{\partial\left(a_{i_1 i_2 \ldots i_k i_{k+1} \ldots i_p} b_{i_{k+1} \ldots i_p}\right)}{\partial a_{j_1 j_2 \ldots j_k j_{k+1} \ldots j_p}} = \delta^{i_1 \ldots i_k i_{k+1} \ldots i_p}_{j_1 \ldots j_k j_{k+1} \ldots j_p} b_{i_{k+1} \ldots i_p} = \delta^{i_1 \ldots i_k}_{j_1 \ldots j_k} b_{j_{k+1} \ldots j_p}, \quad (5)$$

где $$\delta^{i_1 \ldots i_p}_{j_1 \ldots j_p} = \delta^{i_1}_{j_1}\delta^{i_2}_{j_2}\ldots\delta^{i_p}_{j_p} = \begin{cases} 1, \text{ если } i_1 = j_1,\ i_2 = j_2,\ i_p = j_p; \\ 0 \text{ в остальных случаях.} \end{cases}$$

Здесь свертывание производится по индексам $i_{k+1}, \ldots, i_p$. Частная производная — тензор $(k+p)$-й валентности. Например, при $p=2$, $k=1$ получаем $\frac{\partial(a_{i_1 i_2} b_{i_2})}{\partial a_{j_1 j_2}} = \delta^{i_1}_{j_1} b_{j_2}$ — тензор 3-й валентности. Легко проверить, что при дифференцировании по формуле сохраняются обычные правила нахождения частных производных линейной комбинации и произведения координат комитантов.

Формулы дифференцирования комитантов по векторной и матричной переменной, приведенные в [3, 5], можно легко получить, применяя формулу (5), например:

1) $$\frac{\partial(a_{ik} b_{kl})}{\partial b_{ps}} = a_{ik}\delta^{ki}_{ps} = a_{sp}$$

(в матричных обозначениях $\frac{\partial SpAB}{\partial B} = A^{\text{т}}$, где т — знак транспонирования матрицы, $SpAB$ — след матрицы AB).

2) $$\frac{\partial(a_{li} b_{ik} c_{kl})}{\partial b_{ps}} = \delta^{ik}_{ps} a_{li} c_{kl} = a_{lp} c_{sp}$$

(в матричных обозначениях $\frac{\partial SpABC}{\partial B} = A^{\text{т}}C^{\text{т}}$).

3) $$\frac{\partial(x_i a_{ij} x_j)}{\partial x_s} = 2x_i\delta^j_s a_{ij} = 2x_s a_{sj} \text{ при } a_{ij} = a_{ji}$$

(в векторно-матричных обозначениях $\frac{\partial(x^{\text{т}}Ax)}{\partial x} = 2x^{\text{т}}A$).

4) $$\frac{\partial(x_i a_{ij} x_j)}{\partial a_{ps}} = \delta^{ij}_{ps} x_i x_j = x_p x_s$$

(в векторно-матричных обозначениях $\frac{\partial(x^{\text{т}}Ax)}{\partial A} = xx^{\text{т}}$).

5) $$\frac{\partial(a_{ik} a_{kj})}{\partial a_{rs}} = \delta^i_r a_{sj} + a_{ir}\delta^j_s.$$

Формулы частных производных для случая, когда тензоры рассматриваются относительно линейных преобразований переменных,

Таблица 1

Переменная дифференцирования	Дифференцируемая функция			
	Скалярная f	Векторная $\vec{f}=\begin{Bmatrix} f_1 \\ \vdots \\ f_n \end{Bmatrix}$	Матричная $\{f_{ij}\}$	Тензорная $f_{ij\ldots s}$
Скалярная t	Обычная производная $f'(t)=df/dt$ является скалярной функцией (инвариантом)	Векторная функция, вычисленная по формуле $d\vec{f}/dt=\{df_1/dt \ldots df_n/dt\}$	Матричная функция, значениями которой является матрица $\{df_{ij}/dt\}$	Тензорная функция $df_{ij\ldots s}/dt$ (тензор той же валентности, что и дифференцируемая функция)
Векторная $\vec{x}=\begin{Bmatrix} x_1 \\ \vdots \\ x_m \end{Bmatrix}$	Векторная функция, называемая градиентом (вектор-строка), $df/d\vec{x}=\{df/dx_1, \ldots, df/dx_m\}$	Матричная функция, вычисленная по формуле $b_{ij}=\partial f_i/\partial x_j$	Тензорная функция, значениями которой является тензор 3-й валентности $b_{ijk}=\partial f_{ij}/\partial x_k$	Тензорная функция, значениями которой является тензор валентности, на единицу большей, чем $f_{ij\ldots s}$: $b_{ij\ldots sk}=\partial f_{ij\ldots s}/\partial x_k$
Матричная $\{x_{ij}\}$	Матричная функция, значениями которой является матрица $\{b_{ij}\}=\{\partial f/\partial x_{ij}\}$	Тензорная функция, значениями которой является тензор 3-й валентности $b_{ijk}=\partial f_k/\partial x_{ij}$	Тензорная функция, значениями которой является тензор 4-й валентности $b_{jirs}=df_{ij}/dx_{rs}$	Тензорная функция, значениями которой является тензор валентности, на две единицы большей, чем дифференцируемая функция: $b_{ij\ldots srp}=df_{ij\ldots s}/dx_{rp}$
Тензорная $x_{ij\ldots p}$	Тензорная функция, значениями которой является тензор той же валентности, что и переменная дифференцирования, с компонентами $\partial f/\partial a_{ij\ldots p}$	Тензорная функция, значениями которой является тензор валентности, на единицу большей, чем $x_{ij\ldots p}$: $b_{ij\ldots pk}=\partial f_k/\partial x_{ij\ldots p}$	Тензорная функция, значениями которой является тензор валентности, на две единицы большей, чем $x_{ij\ldots p}$: $b_{ijz\ldots p}=\partial f_{ij}/\partial x_{r\ldots p}$	Тензорная функция, значениями которой является тензор суммарной валентности дифференцируемого тензора и тензорной переменной

Таблица 2

Задача	Дано	Найти	Функционал	Градиент функционала	Рекуррентная схема оценивания
1	2	3	4	5	6
1. Идентификация линейных стационарных систем	1) Измерения выхода системы, подлежащей идентификации, z_{tj}; $t=\overline{\tau, \tau+r_t}$, $i=\overline{1, r_z}$ 2) Управление системой u_{ti}, $i=\overline{1, r_u}$ 3) Размерность матрицы параметров системы r_{a*} 4) Матрицы измерения и управления: $D=\{a_{ij}\}$, $i=\overline{1, r_z}$; $j=\overline{1, r_{a*}}$; $B=\{b_{ij}\}$, $i=\overline{1, r_{a*}}$; $j=\overline{1, r_u}$	$\hat{a}_{ij}$ — оценку элементов d^{*}_{ij} матрицы параметров системы $(i, j=\overline{1, r_{a*}})$, доставляющей минимум функционалу I	$I=\|S-z\|^2=(s_{ti}-z_{ti})(s_{ti}-z_{ti})$, где $s_{ti}=d_{ip}y_{pt}$ $(i=\overline{1, r_s}; r_s=r_z; p=\overline{1, r_{a*}})$ — измерения выхода обучающейся модели системы; $y_{tp}=y_{(t-1)r}f_{pr}+b_{ps}u_{ts}$, $f_{pr}=\delta^r_p+ha_{pr}$ — состояния модели и линейное приближение фундаментальной матрицы обучающейся модели системы	$\nabla_{\alpha\beta}=\dfrac{\partial I(a_{\alpha\beta})}{\partial a_{\alpha\beta}}$; $\dfrac{\partial I}{\partial a_{\alpha\beta}}=\dfrac{\partial I}{\partial s_{ti}}\dfrac{\partial s_{tc}}{\partial y_{tp}}\times\dfrac{\partial y_{tp}}{\partial f_{pr}}\dfrac{\partial f_{pr}}{\partial a_{\alpha\beta}}$; $\dfrac{\partial I}{\partial s_{ti}}=2(s_{ti}-z_{ti})$; $\dfrac{\partial s_{ti}}{\partial y_{tp}}=d_{ip}$; $\dfrac{\partial y_{tp}}{\partial f_{pr}}=y_{(t-1)r}+f_{pr}\dfrac{\partial y_{(t-1)r}}{\partial f_{pr}}$; $\dfrac{\partial f_{pr}}{\partial a_{\alpha\beta}}=h\delta^{pr}_{\alpha\beta}$	$\widehat{a}_{ij}=\lim\limits_{k\to\infty} a_{ij}[k]$ 1) Низкий уровень помех — метод наискорейшего спуска: $a_{ij}[k]=a_{ij}[k-1]-\gamma[k]\nabla a_{ij}I[k-1]$, $\gamma[k]$ соответствует принципу наискорейшего спуска $\nabla_{ij}I[k]\nabla_{ij}I[k-1]=0$, возможна лишь поисковая оценка $\gamma[k]$ 2) Низкий уровень помех — метод сопряженных направлений [9]: $a_{ij}[k]=a_{ij}[k-1]+\gamma[k]g_{ij}[k]$, $g_{ij}[k]=\nabla_{ij}I[k]+\dfrac{\|\nabla I[k]\|}{\|\nabla I[k-1]\|}g_{ij}[k-1]$; поисковая оценка $\gamma[k]$ находится из условия $dI[k]/d\gamma[k]=0$ 3) Полная априорная неопределенность в отношении помех — тензорная стохастическая аппроксимация [4]

Таблица 2 (окончание)

1	2	3	4	5	6
2. Аппроксимация сложных систем линейной нестационарной системой	1) Измерения z_{ti} выхода системы, подлежащей аппроксимации 2) Управление системой u_{ti}, $i=\overline{1, r_u}$ 3) Размерность фундаментальной матрицы аппроксимирующей системы r_a и r_q	$\hat{a}_{ijq}$ — тензор параметров линейной нестационарной системы, доставляющей наилучшее приближение наблюдаемого выхода системы Наблюдаемый выход аппроксимирующей системы: $s_{ti} = y_{tp}\alpha_{pi}$; $t=\overline{\tau, \tau+r_t}$; $i=\overline{1, r_s}$, $r_s=r_z$ Состояние системы в момент t: $y_{tp}=f_{tpr}y_{(t-1)r}+ +b_{ps}u_{ts}$ (по t не суммировать) $p,\ r=\overline{1, r_a}$; $s=\overline{1, r_a}$ Линейное приближение фундаментального тензора системы $f_{tpr}=\delta^p_{rt}+ +ha_{prq}l_{qt}$, $p, r=\overline{1, r_a}$, где a_{prq} — разыскиваемый тензор параметров, δ^p_{rt} — единичный тензор	$I=(s_{ti}-z_{ti})(s_{ti}-z_{ti})$	$\nabla_{\alpha\beta\gamma}=\dfrac{\partial I(a_{\alpha\beta\gamma})}{\partial a_{\alpha\beta\gamma}}$, $\alpha,\ \beta=\overline{1, r_a}$, $\gamma=\overline{0, r_q}$; $\nabla_{\alpha\beta\gamma}=\dfrac{\partial I}{\partial s_{ti}}\dfrac{\partial s_{ti}}{\partial y_{tp}}\dfrac{\partial y_{tp}}{\partial f_{tpr}}\dfrac{\partial f_{tpr}}{\partial a_{\alpha\beta\gamma}}$; $\dfrac{\partial y_{tp}}{\partial f_{wuv}}=\delta^{tp}_{wu}y_{(t-1)v}+ +\dfrac{\partial y_{(t-1)r}}{\partial f_{wuv}}f_{tpr}$; $\dfrac{\partial y_{0r}}{\partial f_{wuv}}=0$; $\dfrac{\partial f_{wuv}}{\partial a_{\alpha\beta\gamma}}=h\delta^{uv}_{\alpha\beta}l_{\gamma w}$, δ^{ij}_{kl} — единичный тензор 4-й валентности	$\hat{a}_{ilq}=\lim\limits_{k\to\infty} a_{ijq}[k]$; $a_{ijq}[k]=a_{ijq}[k-1]- -\gamma[k]\nabla_{ijq}I[k]$; варианты выбора $\gamma[k]$ аналогичны задаче 1

3. Описание функции качества системы в пространстве многих факторов формальным степенным рядом	1) Оценки функции качества системы в отдельных точках n-мерного пространства факторов, определяющих качество системы; i — количество точек факторного пространства, в которых получены оценки функции качества 2) Максимальная степень формального ряда, аппроксимирующего функцию качества	3-й валентности, $l_{qt}=l_q[t]$, $q=\overline{0,\ r_q}$ — известные функции времени, например, $l_0[t]=1$; $l_1[t]=ht$, $l_2[t]=h^2t^2,\ldots$ Коэффициенты формального степенного ряда, обеспечивающего наилучшее приближение функции качества системы в n-мерном факторном пространстве $f(x_i)=f_0+a_ix_i+$ $+b_{ij}x_ix_j+\ldots+h_{ij\ldots r}$ $x_i\ldots x_r=A\Pi(x_i)$, где $A=\{f_0,\ a_i,\ b_{ij},\ldots\}$, $\Pi(x_i)=\{1,\ x_i,\ x_ix_j,\ldots,$ $x_ix_j,\ldots,\ x_r\}$	$I=(\varphi_\mu-f_\mu)\ (\varphi_\mu-f_\mu)$, где $f_\mu=f(x_i(\mu))$	$\nabla_A I=\{\nabla_0,\ \nabla_i,\ \nabla_{ij},\ldots,$ $\nabla_{ij\ldots r}\}$; $\nabla_{ij,\ldots,r}=\dfrac{dI}{dh_{ij\ldots r}}=2(f_\mu-$ $-\varphi_\mu)x_i(\mu),\ldots,\ x_r(\mu)$	В о 1 и 2 воз:овое опред(метода спуска: $A[k][k]$

сведены в табл. 1. Аналогичную таблицу можно является общим, универсальным и достаточно обозримым.

Предлагаемый аппарат тензорного исчисления позволяет сформулировать с общих позиций решение таких задач, как:

задачи идентификации, при этом минимизируемым показателем качества процесса идентификации, например, в идентификационном стенде динамических объектов служит функционал, характеризующий близость выходных процессов идентифицируемого объекта и его математической модели на всем обозримом множестве выходных процессов;

задачи аппроксимации функций многих переменных формальными степенными рядами [8] по скалярному критерию близости при степени ряда два и больше, при этом группы коэффициентов аппроксимирующего ряда представляют собой в общем случае тензоры. В такой постановке задача аппроксимации может возникнуть в задачах собственно приближения функций многих переменных, аппроксимации сложных динамических систем по их выходным процессам, распознавания образов, некоторых задач линейной алгебры.

В табл. 2 приведены алгоритмы решения некоторых из упомянутых задач с помощью разработанного аппарата.

ЛИТЕРАТУРА

1. У. Портер. Современные основания общей теории систем. М., «Наука», 1971.
2. Л. Заде, Ч. Дезоер. Теория линейных систем. М., «Наука», 1970.
6. Ф. Р. Гантмахер. Теория матриц. М., «Наука», 1967.
7. А. И. Мальцев. Основы линейной алгебры. М., Гостехиздат, 1956.
8. И. М. Глазман, Ю. И. Любич. Конечномерный линейный анализ. М., «Наука», 1969.
9. К. Спиди, Р. Браун, Д. Гудвин. Теория управления М., «Мир», 1973.

Поступила в редакцию
14 января 1975 г.

УДК 621.391.193

Б. Д. БОРИСОВ, М. И. МОГИЛЬНИЦКИЙ,
А. Г. СЕНИН, М. С. ХАЙРЕТДИНОВ
(Новосибирск)

ВОПРОСЫ ОБУЧЕНИЯ И ПРАКТИЧЕСКОЙ РЕАЛИЗАЦИИ АЛГОРИТМОВ РАСПОЗНАВАНИЯ СЛУЧАЙНЫХ СИГНАЛОВ

Автоматизация процессов контроля во многих ситуациях неразрывно связана с распознаванием временных сигналов. Достаточно ограничиться такими задачами, как акустическая диагностика машин и механизмов, звуковая и ультразвуковая дефектоскопия, обнаружение и классификация шумов, чтобы оценить перспективность этих исследований для задач автоматизации.

Не затрагивая всех вопросов синтеза распознающих устройств, отметим лишь, что выбор признаков и алгоритма обучения, объем памяти и быстродействие — основные факторы, определяющие практическую полезность проектируемых распознающих систем широкого назначения. Анализу этих комплексных вопросов посвящена работа *.

В предлагаемой статье отражены некоторые новые результаты моделирования алгоритмов обучения, положенные в основу разработанного классификатора случайных процессов КСП-170, и его технические особенности, позволяющие оценить возможность приложения для конкретных задач.

Прибор (рис. 1) состоит из трех отдельных блоков: спектрального анализатора, блоков обучения и питания. Структурная схема его представлена на рис. 2.

Случайный процесс $x(t)$ нормируется по амплитуде в усилителе-ограничителе (У — О) так, чтобы уровень сигнала не влиял на распознавание. Нормализованный сигнал подается на 15 параллельных каналов спектрального анализатора (СА), где выделяются текущие энергетические характеристики процесса — признаки, на основе которых

---

* Сенин А. Г. Распознавание случайных сигналов. Новосибирск, «Наука», 1974.