

УДК 681.5.015.42

В.В. Захаров*Севастопольский национальный технический университет**ул. Университетская 33, г. Севастополь, Украина, 99053**E-mail: tk@sevgtu.sebastopol.ua***ВЕКТОРНОЕ ДИФФЕРЕНЦИРОВАНИЕ ПРОИЗВЕДЕНИЙ
МАТРИЧНЫХ ФУНКЦИЙ В ЗАДАЧЕ ПОСТРОЕНИЯ
МОДЕЛИ МОМЕНТОВ МНОГОМЕРНОЙ НЕЛИНЕЙНОЙ
СТОХАСТИЧЕСКОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ**

Предложены обобщения понятий градиента и матрицы Гессе для матричной функции. Определены правила дифференцирования по векторному аргументу классического и прямого матричных произведений функций. На их основе получены соответствующие представления градиента и матрицы Гессе.

Оптимизация производственных процессов на основе методов статистической теории систем управления в пространстве состояний [1] предполагает решение одной из двух задач. Первая состоит в определении плотности распределения вектора состояния стохастической динамической системы, вторая – в определении лишь конечного числа вероятностных моментов, т.е. математического ожидания, корреляционной матрицы и вероятностных моментов высших порядков [2].

Получение уравнений вероятностных моментов нелинейной многомерной дифференциальной системы в негауссовском приближении на основе матричного аналога формулы Ито [3] предполагает аналитическое представление градиента и матрицы Гессе [1] соответствующих матричных функций её вектора состояния. В общем случае это обуславливает необходимость аналитического дифференцирования по векторному аргументу как классического, так и прямого матричных произведений функций [4], и следовательно, определения соответствующих правил векторного дифференцирования. При этом потребность в рассмотрении варианта прямого произведения объясняется уже тем, что само определение произведения матрицы и скалярной величины соответствует определению именно прямого произведения. Очевидно, что для представления матрицы Гессе потребуется рассмотрение двух вариантов правила дифференцирования по векторному аргументу: по вектору-строке и вектору-столбцу.

Целью статьи является, во-первых, рассмотрение двух вариантов определения понятий градиента и матрицы Гессе для матричной функции, а во-вторых – получение соответствующих аналитических представлений градиента и матрицы Гессе для двух типов матричных произведений функций.

Пусть скалярные компоненты $m_{ij}(\mathbf{y})$ ($n_M \times m_M$)-матричной функции $\mathbf{M} = \sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{\mathbf{e}_i \otimes \mathbf{e}_j^T \otimes m_{ij}(\mathbf{y})\}$ ($n_y \times 1$)-векторного аргумента $\mathbf{y}$ дифференцируемы по $\mathbf{y}$ хотя бы один раз. Тогда естественным обобщением понятия градиента матричной функции $\mathbf{M}$ по векторному аргументу $\bar{\mathbf{y}} = \mathbf{y}^T$ представляется вариант его определения через градиенты

$\frac{\partial}{\partial \bar{\mathbf{y}}} m_{ij}(\mathbf{y}) = \sum_{k=1}^{n_y} \mathbf{e}_k^T \otimes \frac{\partial}{\partial y_k} [m_{ij}(\mathbf{y})]$ её скалярных компонентов в виде блочной матрицы

$$\mathbf{M}_{\bar{\mathbf{y}}} = \frac{\partial}{\partial \bar{\mathbf{y}}} [\mathbf{M}] = \sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{\mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \frac{\partial}{\partial \bar{\mathbf{y}}} m_{ij}(\mathbf{y})\}, \quad (1)$$

по векторному аргументу $\mathbf{y}$ – через градиенты $\frac{\partial}{\partial \mathbf{y}} m_{ij}(\mathbf{y}) = \sum_{k=1}^{n_y} \mathbf{e}_k \otimes \frac{\partial}{\partial y_k} [m_{ij}(\mathbf{y})]$ её скалярных компонентов в виде блочной матрицы

$$\mathbf{M}_{\mathbf{y}} = \frac{\partial}{\partial \mathbf{y}} [\mathbf{M}] = \sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{\mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \frac{\partial}{\partial \mathbf{y}} m_{ij}(\mathbf{y})\}, \quad (2)$$

а соответствующим обобщением понятия матрицы Гессе для матричной функции $\mathbf{M}$ – её определение в виде блочной матрицы

$$\mathbf{M}_{yy} = \frac{\partial^2}{\partial \mathbf{y} \partial \mathbf{y}} [\mathbf{M}] = \sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \frac{\partial^2}{\partial y_i \partial y_j} m_{ij}(\mathbf{y}) \}$$

через матрицы Гессе $\frac{\partial^2}{\partial \mathbf{y} \partial \mathbf{y}} m_{ij}(\mathbf{y}) = \sum_{k=1}^{n_y} \sum_{l=1}^{n_y} \mathbf{e}_k \otimes \mathbf{e}_l^T \otimes \frac{\partial^2}{\partial y_k \partial y_l} [m_{ij}(\mathbf{y})]$ скалярных компонентов $\mathbf{M}$, где символ

$\otimes$ означает прямое произведение матриц, T – транспонирование, а $\mathbf{e}_i$, $\mathbf{e}_j$, $\mathbf{e}_k$ и $\mathbf{e}_l$ – единичные векторы [4] соответствующих размеров. При этом в силу определений (1) и (2) правила транспонирования прямого произведения и очевидного равенства $\frac{\partial}{\partial \mathbf{y}} m_{ij}(\mathbf{y}) = \{ \frac{\partial}{\partial y} m_{ij}(\mathbf{y}) \}^T$ связь градиентов по $\bar{\mathbf{y}}$ и $\mathbf{y}$ матричной функции $\mathbf{M}$ определится равенством

$$\begin{aligned} \mathbf{M}_{\bar{\mathbf{y}}} &= \sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes [\frac{\partial}{\partial \mathbf{y}} m_{ij}(\mathbf{y})]^T \} = \\ &= \{ \frac{\partial}{\partial \bar{\mathbf{y}}} [\sum_{i=1}^{n_M} \sum_{j=1}^{m_M} \{ \mathbf{e}_i^T \otimes \mathbf{e}_j \otimes m_{ij}(\mathbf{y}) \}] \}^T = \{ \mathbf{M}_{\bar{\mathbf{y}}}^T \}^T. \end{aligned} \quad (3)$$

Метод описания градиентов по $\bar{\mathbf{y}}$ матричных произведений основывается на двух теоремах и нескольких следствиях из них. Они определяют два важных и отнюдь не очевидных свойства операций произведения и прямого произведения матриц.

Теорема 1. Для любых матриц $\mathbf{M}_{n \times m}$, $\mathbf{M}_{k \times l}$, $\mathbf{M}_{m \times p}$, $\mathbf{M}_{l \times q}$ и любых, определяющих их размеры, целых положительных n, m, l, k, p, q справедливо равенство

$$(\mathbf{M}_{n \times m} \mathbf{M}_{m \times p}) \otimes (\mathbf{M}_{k \times l} \mathbf{M}_{l \times q}) = (\mathbf{M}_{n \times m} \otimes \mathbf{M}_{k \times l}) (\mathbf{M}_{m \times p} \otimes \mathbf{M}_{l \times q}). \quad (4)$$

Доказательство теоремы 1 не вызывает затруднений и аналогично доказательству, приведённому в [4] для квадратных матриц.

Следствие 1.1. Пусть в (4) $\mathbf{M}_{l \times q} = \mathbf{I}_l$ и $n = m = 1$. Тогда

$$\mathbf{M}_{k \times l} (\mathbf{M}_{1 \times p} \otimes \mathbf{I}_l) = \mathbf{M}_{1 \times p} \otimes \mathbf{M}_{k \times l},$$

где здесь и далее $\mathbf{I}_l$ – единичная $(l \times l)$ -матрица.

Следствие 1.2. Транспонируя соотношение следствия 1.1, получим

$$(\mathbf{M}_{p \times l} \otimes \mathbf{I}_l) \mathbf{M}_{l \times k} = \mathbf{M}_{p \times l} \otimes \mathbf{M}_{l \times k}.$$

Следствие 1.3. Если в (4) $q = n = m = 1$, то

$$\mathbf{M}_{1 \times p} \otimes (\mathbf{M}_{k \times l} \mathbf{M}_{l \times 1}) = \mathbf{M}_{k \times l} (\mathbf{M}_{1 \times p} \otimes \mathbf{M}_{l \times 1}).$$

Теорема 2. Для любых $(n_\Phi \times m_\Phi)$ -матрицы Φ и $(n_\Psi \times m_\Psi)$ -матрицы Ψ существуют такие $(n_\Phi \times n_\Psi) \times (n_\Phi \times n_\Psi)$ -матрица перестановок строк

$$\mathbf{P}_{n_\Phi \times n_\Psi} = \sum_{i=1}^{n_\Phi} \{ \mathbf{e}_i \otimes \mathbf{I}_{n_\Psi} \otimes \mathbf{e}_i^T \}$$

и $(m_\Phi \times m_\Psi) \times (m_\Phi \times m_\Psi)$ -матрица перестановок столбцов

$$\mathbf{P}_{m_\Phi \times m_\Psi} = \sum_{i=1}^{m_\Phi} \{ \mathbf{e}_i^T \otimes \mathbf{I}_{m_\Psi} \otimes \mathbf{e}_i \},$$

что

$$\Phi \otimes \Psi = \mathbf{P}_{n_\Phi \times n_\Psi} (\Psi \otimes \Phi) \mathbf{P}_{m_\Phi \times m_\Psi}. \quad (5)$$

Доказательство. Представим $(n_\Phi \times m_\Phi)$ -матрицу Φ в виде

$$\Phi = \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \phi_{ij} \},$$

её i, j -тую компоненту как $\phi_{ij} = \mathbf{e}_i^T \Phi \mathbf{e}_j$, а $(n_\Psi \times m_\Psi)$ -матрицу Ψ как $\Psi = \mathbf{I}_{n_\Psi} \Psi \mathbf{I}_{m_\Psi}$. В новых обозначениях прямое произведение в (5) принимает вид

$$\Phi \otimes \Psi = \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes (\mathbf{I}_{n_\Psi} \Psi \mathbf{I}_{m_\Psi}) \otimes (\mathbf{e}_i^T \Phi \mathbf{e}_j) \}. \quad (6)$$

Легко увидеть, что на основании теоремы 1 получаем:

$$(\mathbf{I}_{n_\Psi} \Psi \mathbf{I}_{m_\Psi}) \otimes (\mathbf{e}_i^T \Phi \mathbf{e}_j) = (\mathbf{I}_{n_\Psi} \otimes \mathbf{e}_i^T) (\Psi \otimes \Phi) (\mathbf{I}_{m_\Psi} \otimes \mathbf{e}_j). \quad (7)$$

Тогда, подставив (7) в (6) и воспользовавшись для преобразования ещё раз теоремой 1, окончательно получим (5).

Заметим, что с помощью теоремы 2 легко получить результат, приведённый в [4], устанавливающий характер связи между матрицами перестановок строк и столбцов, коммутирующих прямое произведение квадратных матриц. Так, если $n_\Phi = m_\Phi$ и $n_\Psi = m_\Psi$, матрицы перестановок удовлетворяют равенству $\mathbf{P}_{n_\Phi \times n_\Psi} = \mathbf{P}_{m_\Phi \times m_\Psi}^T$. Однако, в общем случае равенство не выполняется.

Сформулируем два полезных следствия из теоремы 2.

Следствие 2.1. Пусть хотя бы одна из матриц Φ или Ψ в (5) – вектор. Тогда, если:

- 1) $(m_\Psi = 1) \cup (m_\Phi = 1)$, то $\Phi \otimes \Psi = \mathbf{P}_{n_\Phi \times n_\Psi} (\Psi \otimes \Phi)$;
- 2) $(n_\Psi = 1) \cup (n_\Phi = 1)$, то $\Phi \otimes \Psi = (\Psi \otimes \Phi) \mathbf{P}_{m_\Phi \times m_\Psi}$;
- 3) $(m_\Phi = n_\Psi = 1) \cup (n_\Phi = m_\Psi = 1)$, то $\Phi \otimes \Psi = \Psi \otimes \Phi$.

Это означает, что прямое произведение столбца и строки (либо строки и столбца) коммутативно [4].

Следствие 2.2. Легко показать, что матрицы перестановок в (5) такие, что

$$\mathbf{P}_{n_\Phi \times n_\Psi} = \mathbf{P}_{n_\Psi \times n_\Phi}^T \quad \text{и} \quad \mathbf{P}_{m_\Phi \times m_\Psi} = \mathbf{P}_{m_\Psi \times m_\Phi}^T.$$

Пусть $\Phi \Psi$ – произведение $(n_\Phi \times m_\Phi)$ -матричной Φ и $(n_\Psi \times m_\Psi)$ -матричной Ψ функций $(n_y \times 1)$ -векторного аргумента $\mathbf{y}$, дифференцируемых по $\mathbf{y}$ хотя бы один раз. Очевидно [4], что $m_\Phi = n_\Psi$. По определению (1), градиент по $\bar{\mathbf{y}}$ матричной функции $\Phi \Psi$ – есть блочная матрица

$$\frac{\partial}{\partial \bar{\mathbf{y}}} [\Phi \Psi] = \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \frac{\partial}{\partial \bar{\mathbf{y}}} [\mathbf{e}_i^T \Phi \Psi \mathbf{e}_j] \} \quad (8)$$

градиентов $\frac{\partial}{\partial \bar{\mathbf{y}}} [\mathbf{e}_i^T \Phi \Psi \mathbf{e}_j]$ скалярных компонентов $\mathbf{e}_i^T \Phi \Psi \mathbf{e}_j$, причём $\mathbf{e}_i^T \Phi \Psi \mathbf{e}_j = \sum_{k=1}^{m_\Phi} \{ \phi_{ik} \psi_{kj} \}$. Тогда в силу определения градиента произведения скалярных функций получаем:

$$\frac{\partial}{\partial \bar{\mathbf{y}}} [\mathbf{e}_i^T \Phi \Psi \mathbf{e}_j] = \sum_{k=1}^{m_\Phi} \{ \frac{\partial}{\partial \bar{\mathbf{y}}} [\phi_{ik}] \otimes \psi_{kj} + \phi_{ik} \otimes \frac{\partial}{\partial \bar{\mathbf{y}}} [\psi_{kj}] \}. \quad (9)$$

Перейдём в (9) к матричному описанию скалярных компонентов функций Φ , Ψ и их градиентов: $\phi_{ik} = \mathbf{e}_i^T \Phi \mathbf{e}_k$, $\psi_{kj} = \mathbf{e}_k^T \Psi \mathbf{e}_j$, $\frac{\partial}{\partial \bar{\mathbf{y}}} [\phi_{ik}] = \mathbf{e}_i^T \Phi_{\bar{\mathbf{y}}} (\mathbf{e}_k \otimes \mathbf{I}_{n_y})$ и $\frac{\partial}{\partial \bar{\mathbf{y}}} [\psi_{kj}] = \mathbf{e}_k^T \Psi_{\bar{\mathbf{y}}} (\mathbf{e}_j \otimes \mathbf{I}_{n_y})$. В силу теоремы 1 и следствия 2.1 (пункт 1)

$$\sum_{k=1}^{m_\Phi} \{ \frac{\partial}{\partial \bar{\mathbf{y}}} [\phi_{ik}] \otimes \psi_{kj} \} = \mathbf{e}_i^T \Phi_{\bar{\mathbf{y}}} (\Psi \otimes \mathbf{I}_{n_y}) (\mathbf{e}_j \otimes \mathbf{I}_{n_y}),$$

а в силу следствия 1.3 и очевидного равенства $\sum_{k=1}^{m_\Phi} \{ \mathbf{e}_k \otimes \mathbf{e}_k^T \} = \mathbf{I}_{m_\Phi}$ можно записать:

$$\sum_{k=1}^{m_\Phi} \{ \phi_{ik} \otimes \frac{\partial}{\partial \bar{\mathbf{y}}} [\psi_{kj}] \} = \mathbf{e}_i^T \Phi \Psi_{\bar{\mathbf{y}}} (\mathbf{e}_j \otimes \mathbf{I}_{n_y}).$$

Окончательно с учётом полученных представлений слагаемых в правой части (9), (8) принимает вид

$$\frac{\partial}{\partial \bar{\mathbf{y}}} [\Phi \Psi] = \Phi_{\bar{\mathbf{y}}} (\Psi \otimes \mathbf{I}_{n_y}) + \Phi \Psi_{\bar{\mathbf{y}}}. \quad (10)$$

Из (3) следует, что если $\mathbf{M} = \Phi \Psi$ ($\mathbf{M}^T = \Psi^T \Phi^T$), то представление градиента $\frac{\partial}{\partial \mathbf{y}}[\Phi \Psi]$ получается после формальной замены в (10) Φ , Ψ на Ψ^T , Φ^T соответственно и транспонирования полученной таким образом модификации в виде

$$\frac{\partial}{\partial \mathbf{y}}[\Phi \Psi] = \Phi_{\bar{\mathbf{y}}} \Psi + (\Phi \otimes \mathbf{I}_{n_y}) \Psi_{\mathbf{y}}. \quad (11)$$

Пусть теперь $\Phi \otimes \Psi$ – прямое произведение $(n_\Phi \times m_\Phi)$ -матричной Φ и $(n_\Psi \times m_\Psi)$ -матричной Ψ функций векторного аргумента $\mathbf{y}$, также дифференцируемых по $\mathbf{y}$ хотя бы один раз. По определению (1), градиент по $\bar{\mathbf{y}}$ $(n_\Phi \times n_\Psi) \times (m_\Psi \times m_\Phi)$ -матричной функции $\Phi \otimes \Psi$ – суть блочная матрица

$$\frac{\partial}{\partial \bar{\mathbf{y}}}[\Phi \otimes \Psi] = \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \sum_{k=1}^{n_\Psi} \sum_{l=1}^{m_\Psi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \mathbf{e}_k \otimes \mathbf{e}_l^T \otimes \frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij} \psi_{kl}] \} \quad (12)$$

градиентов $\frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij} \psi_{kl}]$ скалярных компонентов $\phi_{ij} \psi_{kl}$, причём

$$\frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij} \psi_{kl}] = \frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij}] \psi_{kl} + \phi_{ij} \frac{\partial}{\partial \bar{\mathbf{y}}}[\psi_{kl}].$$

В силу следствия 2.1 (пункт 2) и теоремы 1 имеем

$$\sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \sum_{k=1}^{n_\Psi} \sum_{l=1}^{m_\Psi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \mathbf{e}_k \otimes \mathbf{e}_l^T \otimes (\phi_{ij} \frac{\partial}{\partial \bar{\mathbf{y}}}[\psi_{kl}]) \} = \Phi \otimes \Psi_{\bar{\mathbf{y}}}$$

и

$$\begin{aligned} & \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \sum_{k=1}^{n_\Psi} \sum_{l=1}^{m_\Psi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \mathbf{e}_k \otimes \mathbf{e}_l^T \otimes (\frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij}] \psi_{kl}) \} = \\ & = \sum_{i=1}^{n_\Phi} \sum_{j=1}^{m_\Phi} \{ \mathbf{e}_i \otimes \mathbf{e}_j^T \otimes \Psi \otimes \frac{\partial}{\partial \bar{\mathbf{y}}}[\phi_{ij}] \} = (\Phi_{\bar{\mathbf{y}}} \otimes \Psi)(\mathbf{I}_{m_\Phi} \otimes \mathbf{P}_{m_\Psi \times n_y}). \end{aligned}$$

Тогда с учётом представлений слагаемых в правой части (12) принимает вид

$$\frac{\partial}{\partial \bar{\mathbf{y}}}[\Phi \otimes \Psi] = (\Phi_{\bar{\mathbf{y}}} \otimes \Psi)(\mathbf{I}_{m_\Phi} \otimes \mathbf{P}_{m_\Psi \times n_y}) + \Phi \otimes \Psi_{\bar{\mathbf{y}}}, \quad (13)$$

где $\mathbf{P}_{m_\Psi \times n_y} = \sum_{i=1}^{m_\Psi} \{ \mathbf{e}_i^T \otimes \mathbf{I}_{n_y} \otimes \mathbf{e}_i \}$.

Из (3) следует, что если $\mathbf{M} = \Phi \otimes \Psi$ ($\mathbf{M}^T = \Phi^T \otimes \Psi^T$), то представление градиента $\frac{\partial}{\partial \mathbf{y}}[\Phi \otimes \Psi]$ получается после формальной замены в (13) Φ , Ψ , m_Φ , m_Ψ на Φ^T , Ψ^T , n_Φ , n_Ψ соответственно и транспонирования полученной таким образом модификации в виде

$$\frac{\partial}{\partial \mathbf{y}}[\Phi \otimes \Psi] = (\mathbf{I}_{n_\Phi} \otimes \mathbf{P}_{n_\Psi \times n_y}^T)(\Phi_{\mathbf{y}} \otimes \Psi) + \Phi \otimes \Psi_{\mathbf{y}}, \quad (14)$$

где $\mathbf{P}_{n_\Psi \times n_y}^T = \sum_{i=1}^{n_y} \{ \mathbf{e}_i^T \otimes \mathbf{I}_{n_\Psi} \otimes \mathbf{e}_i \}$.

На основе (10), (11) и (13), (14) легко получаются и соответствующие двум вариантам матричных произведений искомые представления матрицы Гессе.

Для получения представления матрицы Гессе произведения $\Phi \Psi$ продифференцируем (10) по $\mathbf{y}$. Первое слагаемое полученного в результате представления $\frac{\partial^2}{\partial \mathbf{y} \partial \bar{\mathbf{y}}}[\Phi \Psi]$ преобразуется на основе (11) и (14) к виду

$$\begin{aligned} & \frac{\partial}{\partial \mathbf{y}}\{ \Phi_{\bar{\mathbf{y}}}(\Psi \otimes \mathbf{I}_{n_y}) \} = \Phi_{\mathbf{y}\mathbf{y}}(\Psi \otimes \mathbf{I}_{n_y}) + (\Phi_{\bar{\mathbf{y}}} \otimes \mathbf{I}_{n_y}) \frac{\partial}{\partial \mathbf{y}}[\Psi \otimes \mathbf{I}_{n_y}] = \\ & = \Phi_{\mathbf{y}\mathbf{y}}(\Psi \otimes \mathbf{I}_{n_y}) + (\Phi_{\bar{\mathbf{y}}} \otimes \mathbf{I}_{n_y})(\mathbf{I}_{n_\Psi} \otimes \mathbf{P}_{n_y \times n_y})(\Psi_{\mathbf{y}} \otimes \mathbf{I}_{n_y}), \end{aligned}$$

а второе – на основе (11) к виду

$$\frac{\partial}{\partial \mathbf{y}}\{ \Phi \Psi_{\bar{\mathbf{y}}} \} = \Phi_{\mathbf{y}} \Psi_{\bar{\mathbf{y}}} + (\Phi \otimes \mathbf{I}_{n_y}) \Psi_{\mathbf{y}\mathbf{y}}.$$

Окончательно с учётом полученных представлений слагаемых искомое выражение принимает вид

$$\begin{aligned} \frac{\partial^2}{\partial y \partial \bar{y}} [\Phi \Psi] = & \Phi_{yy} (\Psi \otimes I_{n_y}) + (\Phi \otimes I_{n_y}) \Psi_{yy} + \Phi_y \Psi_{\bar{y}} + \\ & + (\Phi_{\bar{y}} \otimes I_{n_y}) (I_{n_\Psi} \otimes P_{n_y \times n_y}) (\Psi_y \otimes I_{n_y}), \end{aligned} \quad (15)$$

где $P_{n_y \times n_y} = \sum_{i=1}^{n_y} \{e_i^T \otimes I_{n_y} \otimes e_i\}$.

Представление матрицы Гессе прямого произведения $\Phi \otimes \Psi$ получается аналогично. Слагаемые получаемого в результате дифференцирования (13) по y представления $\frac{\partial^2}{\partial y \partial \bar{y}} [\Phi \otimes \Psi]$ преобразуются на основе (14) к виду

$$\frac{\partial}{\partial y} [\Phi_{\bar{y}} \otimes \Psi] = (I_{n_\Phi} \otimes P_{n_y \times n_\Psi}) (\Phi_{yy} \otimes \Psi) + \Phi_{\bar{y}} \otimes \Psi_y$$

и

$$\frac{\partial}{\partial y} \{\Phi \otimes \Psi_{\bar{y}}\} = (I_{n_\Phi} \otimes P_{n_y \times n_\Psi}) (\Phi_y \otimes \Psi_{\bar{y}}) + \Phi \otimes \Psi_{yy}$$

соответственно. С учётом полученных представлений слагаемых, искомое для матрицы Гессе принимает вид

$$\begin{aligned} \frac{\partial^2}{\partial y \partial \bar{y}} [\Phi \otimes \Psi] = & (I_{n_\Phi} \otimes P_{n_y \times n_\Psi}) (\Phi_{yy} \otimes \Psi) (I_{m_\Phi} \otimes P_{m_\Psi \times n_y}) + \\ & + (\Phi_{\bar{y}} \otimes \Psi_y) (I_{m_\Phi} \otimes P_{m_\Psi \times n_y}) + \\ & + (I_{n_\Phi} \otimes P_{n_y \times n_\Psi}) (\Phi_y \otimes \Psi_{\bar{y}}) + \Phi \otimes \Psi_{yy}. \end{aligned} \quad (16)$$

Если $m_\Phi = n_\Psi = 1$, то классические и прямое матричные произведения оказываются тождественными, причём $\Phi \Psi = \Phi \otimes \Psi$ – векторное произведение [4]. Легко убедиться, что в этом случае градиенты (10) и (13) в силу следствий 1.1 и 2.1 (пункт 2) совпадают и равны

$$\frac{\partial}{\partial y} [\Phi \otimes \Psi] = \Psi \otimes \Phi_{\bar{y}} + \Phi \otimes \Psi_{\bar{y}},$$

а матрицы Гессе (15) и (16) в силу следствий 1.1, 1.2 и 2.1 (пункт 2) принимают вид

$$\frac{\partial^2}{\partial y \partial \bar{y}} [\Phi \otimes \Psi] = \Psi \otimes \Phi_{yy} + \Phi \otimes \Psi_{yy} + \Phi_y \otimes \Psi_{\bar{y}} + (\Phi_{\bar{y}} \otimes \Psi_y) P_{m_\Psi \times n_y}.$$

Другим известным [5] вариантом обобщения понятия градиента для $(n_M \times m_M)$ -матричной функции M по $(1 \times n_y)$ - и $(n_y \times 1)$ -векторным аргументам $\bar{y}$ и y являются его определения в виде блочных матриц

$$\nabla_{\bar{y}} M = \frac{\partial}{\partial \bar{y}} [M] = \sum_{k=1}^{n_y} \{e_k^T \otimes \frac{\partial}{\partial y_k} [M]\} \quad (17)$$

и

$$\nabla_y M = \frac{\partial}{\partial y} [M] = \sum_{k=1}^{n_y} \{e_k \otimes \frac{\partial}{\partial y_k} [M]\}, \quad (18)$$

а соответствующим обобщением понятия её матрицы Гессе – в виде

$$\nabla_y^2 M = \frac{\partial^2}{\partial y \partial \bar{y}} [M] = \sum_{i=1}^{n_y} \sum_{j=1}^{n_y} \{e_i \otimes e_j^T \otimes \frac{\partial^2}{\partial y_i \partial y_j} [M]\}. \quad (19)$$

Очевидно, что в этом случае искомые представления градиентов $\nabla_{\bar{y}} M$, $\nabla_y M$ и матрицы Гессе $\nabla_y^2 M$ матричного произведения $M = \Phi \Psi$ легко удастся получить из их аналогов (10), (11) и (15) на основе соотношений, связывающих $\nabla_{\bar{y}} M$, $\nabla_y M$ и $\nabla_y^2 M$ с $M_{\bar{y}}$, M_y и M_{yy} соответственно. Так из (17) в силу (1) и следствия 2.1 (пункт 2) следует, что

$$\nabla_{\bar{y}} M = M_{\bar{y}} P_{n_y \times m_M}, \quad (20)$$

из (18) в силу (2) и следствия 2.1 (пункт 1)

$$\nabla_y M = P_{n_y \times n_M} M_y, \quad (21)$$

а из (19) в силу (20), (21) и (11)

$$\nabla_y^2 \mathbf{M} = \mathbf{P}_{n_y \times n_M} \mathbf{M}_{yy} \mathbf{P}_{n_y \times m_M}.$$

Здесь $\mathbf{P}_{n_y \times m_M} = \sum_{i=1}^{n_y} \{\mathbf{e}_i^T \otimes \mathbf{I}_{m_M} \otimes \mathbf{e}_i\}$, $\mathbf{P}_{n_y \times n_M} = \sum_{i=1}^{n_y} \{\mathbf{e}_i \otimes \mathbf{I}_{n_M} \otimes \mathbf{e}_i^T\}$ – матрицы перестановок.

Следует отметить взаимосвязь двух рассмотренных матричных операций умножения, очевидную как из приведённых теорем, так и из полученных представлений градиентов (10), (11), (13), (14) и матриц Гессе (15), (16) для матричных произведений функций. При этом заметим, что уже при описании правил (10), (11) векторного дифференцирования произведения матриц возникает потребность в привлечении операции прямого произведения. Известный же альтернативный подход к описанию правила векторного дифференцирования произведения матриц на основе определения дополнительной нестандартной операции матричного умножения [5] представляется нам не целесообразным, т.к. предполагает разработку соответствующей матричной алгебры и, очевидно, приводит к возникновению проблем при применении полученных таким способом аналитических представлений.

Полученные представления градиента (13) и матрицы Гессе (15) являются необходимым инструментом для продолжения исследований. Их направление связано с проблематикой построения модели моментов нелинейной стохастической дифференциальной системы в негауссовском приближении и параметрической оптимизацией дифференциальной системы на основе градиентных методов. В рамках этого направления на основе формулы Ито будут получены матричные уравнения центральных моментов до четвёртого порядка включительно в негауссовском приближении и уравнение чувствительности для градиента дисперсионной матрицы вектора состояния стохастической дифференциальной системы по оптимизируемым параметрам. При этом в качестве негауссовского будет выбрано приближение на основе аппроксимации плотности распределения конечным отрезком разложения по многомерным полиномам Эрмита. В более отдалённой перспективе представляется целесообразным исследование проблематики эффективной алгоритмизации метода корреляционного анализа в негауссовском приближении и решения задачи оценки погрешности корреляционного анализа в гауссовском приближении нелинейной стохастической дифференциальной системы на основе её матричной модели моментов.

В качестве примера применения предложенных представлений (10), (13), (15), (16) приведём полученное в [3] на основе матричного аналога формулы Ито уравнение для матрицы центральных моментов третьего порядка $(n_y \times 1)$ -вектора $\mathbf{y}(t)$ – суть решения стохастического нелинейного дифференциального уравнения Ито [2] вида

$$d\mathbf{y}(t) = \mathbf{r}[\mathbf{y}(t), t]dt + \mathbf{C}(t)d\boldsymbol{\omega}(t), \quad (22)$$

где $\mathbf{r}[\mathbf{y}(t), t] = \mathbf{A}(t)\mathbf{y}(t) + \mathbf{B}(t)\mathbf{f}[\mathbf{y}(t), t]$. В этом случае возникает необходимость аналитического представления градиента $\mathbf{G}_y^k[\mathbf{y}]$ и матрицы Гессе $\mathbf{G}_{yy}^k[\mathbf{y}]$ $(n_y \times n_y)$ -матричной функции $\mathbf{G}^k[\mathbf{y}] = \mathbf{G}[\mathbf{y}] \otimes \mathbf{y}_k$, где $\mathbf{G}[\mathbf{y}] = \mathbf{y} \otimes \mathbf{y}^T$ – матрица Грама [4]. В силу $m_\Phi = n_\Psi = 1$, для получения представлений $\mathbf{G}_y^k[\mathbf{y}]$ и $\mathbf{G}_{yy}^k[\mathbf{y}]$ целесообразно воспользоваться вариантами формул для градиента (10), (13) и матрицы Гессе (15), (16) матричных произведений, соответствующими частному случаю – векторному произведению функций. В результате искомые представления определяются как

$$\mathbf{G}_y^k[\mathbf{y}] = \mathbf{G}_y[\mathbf{y}] \otimes \mathbf{y}_k + \mathbf{G}[\mathbf{y}] \otimes \mathbf{e}_k^T, \quad (23)$$

$$\mathbf{G}_{yy}^k[\mathbf{y}] = \mathbf{G}_{yy}[\mathbf{y}] \otimes \mathbf{y}_k + \mathbf{G}_y[\mathbf{y}] \otimes \mathbf{e}_k + \{\mathbf{G}_y[\mathbf{y}]\}^T \otimes \mathbf{e}_k^T, \quad (24)$$

где $\mathbf{G}_y[\mathbf{y}] = \mathbf{y}^T \otimes \mathbf{I}_n + \mathbf{y} \otimes \mathbf{e}^T$, $\mathbf{G}_{yy}[\mathbf{y}] = \mathbf{P}_{n \times n} + \mathbf{e} \otimes \mathbf{e}^T$, $\mathbf{P}_{n \times n} = \sum_{i=1}^n \{\mathbf{e}_i \otimes \mathbf{I}_n \otimes \mathbf{e}_i^T\}$, а $\mathbf{y}_k = \mathbf{e}_k^T \mathbf{y}$. Окончательно,

с учётом (23) и (24), формула Ито [3] для матричной функции $\mathbf{G}^k[\mathbf{y}]$ принимает вид

$$d\mathbf{G}^k[\mathbf{y}] = \mathbf{R}_k[\mathbf{y}, t] + \mathbf{R}_k^T[\mathbf{y}, t] + \sum_{i=1}^n \{\mathbf{e}_i \otimes (\mathbf{e}_k^T \mathbf{R}_i[\mathbf{y}, t])\}, \quad (25)$$

а соответствующее (25) искомое уравнение k -ого блока $\mathbf{M}_k = E\{\mathbf{G}^k[\mathbf{y}]\}$, $k = \overline{1, n}$ блочной матрицы моментов третьего порядка $\sum_{k=1}^n \{\mathbf{e}_k^T \otimes \mathbf{M}_k\}$ решения стохастического нелинейного дифференциального уравнения (22) определится как

$$\frac{d}{dt} \mathbf{M}_k = \bar{\mathbf{R}}_k + \bar{\mathbf{R}}_k^T + \sum_{i=1}^n \{\mathbf{e}_i \otimes (\mathbf{e}_k^T \bar{\mathbf{R}}_i)\},$$

где $\mathbf{R}_j[\mathbf{y}, t] = \mathbf{R}[\mathbf{y}, t] \otimes y_j$, $\bar{\mathbf{R}}_j = E\{\mathbf{R}_j[\mathbf{y}, t]\}$, $j \in \{k, i\}$, $\mathbf{R}[\mathbf{y}, t] = \mathbf{y} \otimes \mathbf{r}^T[\mathbf{y}, t]$.

В заключение заметим, что подобным же образом на основе формул для градиента (10), (13), матрицы Гессе (15), (16) и матричного аналога формулы Ито можно получить в матричном виде систему уравнений вероятностных моментов любого порядка компонент решения уравнения вида (22). Практика же применения таких моделей для статистического анализа и синтеза стохастических дифференциальных систем, в частности АСУ ТП, свидетельствует о том, что в большинстве случаев при построении модели моментов при негауссовском приближении плотности вероятности решения уравнения (22) оказывается оправданным ограничение лишь моментами до четвёртого порядка включительно. При этом величины вероятностных моментов третьего порядка характеризуют асимметрию, а четвёртого порядка – «плосковершинность» или «островершинность» кривой негауссовской плотности вероятности по сравнению с гауссовской [2].

Библиографический список

1. Справочник по теории автоматического управления / Под ред. А.А. Красовского. — М.: Наука, 1987. — 712 с.
2. Пугачёв В.С. Теория стохастических систем: учеб. пособие / В.С. Пугачёв, И.Н. Сеницын. — М.: Изд-во "Логос Пресс", 2005. — 999 с.
3. Захаров В.В. Матричный аналог формулы Ито и его приложение к корреляционному анализу / В.В. Захаров // Вестник СевГТУ. Сер. Информатика, электроника, связь: сб. науч. тр. — Севастополь, 2007. — Вып. 82. — С. 27–32.
4. Ланкастер П. Теория матриц / П. Ланкастер. — М.: Наука, 1978. — 280 с.
5. Хофер Э. Численные методы оптимизации / Э. Хофер, Р. Лундерштедт. — М.: Машиностроение, 1981. — 192 с.

Поступила в редакцию 22.06.2007 г.