Формулы дифференцирования следа матриц, включающих обратные матрицы

Аппарат дифференцирования следа матриц по матрицам (ДСМ) обеспечивает предельный лаконизм соответствующих преобразований, сопровождающих решение многих экстремальных задач, связанных с анализом необходимых и достаточных условий отыскания экстремума функций многих переменных. В таких задачах, как, например, определение параметров полиномиальной регрессии методом наименьших квадратов, формирования алгоритмов оптимальной и субоптимальной фильтрации при решении задач оценки векторов состояний динамических систем и многих других задачах, применение аппарата ДСМ исключает необходимость выписывать отдельные скалярные уравнения для каждого искомого параметра и решать полученную систему уравнений, что резко облегчает анализ, обеспечивая возможность получения результата сразу в векторно-матричной форме.

Наибольшее распространение получили формулы ДСМ, в которых дифференцируемые по матрице **X** выражения включали след линейных преобразований различных матриц согласованной размерности, например, **AXB, AXBX,** $\mathbf{AX}^{T}\mathbf{B}$. Эти формулы и основные свойства следа матрицы приведены в >>>>.

В задачах анализа чувствительности динамических систем в операторной форме возникает необходимость расширения спектра этих выражений путем включения обратных матриц в матричные выражения[1].

Ниже приводятся результаты, полученные автором для этих выражений (след матрицы обозначен символом ***tr*, E** – единичная матрица).

Общая формула, используемая при этом, имеет вид:

$$\frac{\partial}{\partial \mathbf{A}} tr\{[\mathbf{D}(\mathbf{A})]^{-1}\} = \frac{\partial}{\partial \mathbf{D}(\mathbf{A})} tr\{[\mathbf{D}(\mathbf{A})]^{-1}\}\cdot \frac{\partial}{\partial \mathbf{A}} tr\{[\mathbf{D}(\mathbf{A})]\}. \qquad (1)$$

1. $\frac{\partial}{\partial \mathbf{A}} tr[\mathbf{A}^{-1}] = -(\mathbf{A}^{-2})^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[\mathbf{A}^{-1}] = -(\mathbf{A}^{-2})$.

2. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{A})^{-1}] = -[(\mathbf{E}+\mathbf{A})^{-2}]^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[\mathbf{A}^{-1}] = -(\mathbf{E}+\mathbf{A})^{-2}$.

3. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{A})^{-1}\mathbf{A}] = [(\mathbf{E}+\mathbf{A})^{-2}]^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{E}+\mathbf{A})^{-1}\mathbf{A}] = (\mathbf{E}+\mathbf{A})^{-2}$.

4. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{AB})^{-1}] = -[(\mathbf{AB})^{-2}]^{T}\cdot\mathbf{B}^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{AB})^{-1}] = -\mathbf{B}\cdot(\mathbf{AB})^{-2}$.

5. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{AB})^{-1}] = -[(\mathbf{E}+\mathbf{AB})^{-2}]^{T}\cdot\mathbf{B}^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{E}+\mathbf{AB})^{-1}] = - -\mathbf{B}\cdot(\mathbf{E}+\mathbf{AB})^{-2}$.

6. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{AB})^{-1}\mathbf{AB}] = [(\mathbf{E}+\mathbf{AB})^{-2}]^{T}\cdot\mathbf{B}^{T}$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{E}+\mathbf{AB})^{-1}\mathbf{AB}] = \mathbf{B}\cdot(\mathbf{E}+\mathbf{AB})^{-2}$.

7. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{BA})^{-1}\mathbf{BA}] = \mathbf{B}^{T}[(\mathbf{E}+\mathbf{BA})^{-2}]^{T}\cdot$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{E}+\mathbf{BA})^{-1}\mathbf{BA}] = (\mathbf{E}+\mathbf{BA})^{-2}\cdot\mathbf{B}$

8. $\frac{\partial}{\partial \mathbf{A}} tr[(\mathbf{E}+\mathbf{BA})^{-1}] = -\mathbf{B}^{T}[(\mathbf{E}+\mathbf{BA})^{-2}]^{T}\cdot$; $\frac{\partial}{\partial \mathbf{A}^{T}} tr[(\mathbf{E}+\mathbf{BA})^{-1}] = -(\mathbf{E}+\mathbf{BA})^{-2}\cdot\mathbf{B}$

В разделе **&&&&** приводятся символьные и числовые проверки этих формул.

1. 1 Ивановский Р.И. Проблемы чувствительности в задачах моделирования, обработки информации и управления // Гироскопия и навигация № 1 (72), 2011. с. 90–104.