www.esa-conference.ru

О вычислении целых функций от матриц по Леверье – Фаддееву

Илюшечкин Никита Васильевич, инженер
АО «Концерн «Моринформсистема – Агат», г. Москва

Предлагается алгоритм вычисления целой функции от матрицы, не требующий предварительного определения её собственных значений. Матричная функция вычисляется как многочлен степени на единицу меньшей порядка матрицы. Коэффициенты этого многочлена накапливаются посредством простой процедуры, учитывающей коэффициенты степенного ряда для исходной целой функции.

Ключевые слова: *функция от матрицы, целая функция, многочлен Лагранжа, полный симметрический многочлен.*

Определение функции от матрицы, определённой на спектре этой матрицы, приводится в [1,с.96]. Для нас имеет значение, что, если Z – матрица порядка n, то функция g от матрицы Z представима многочленом H_g степени не более $n-1$ (интерполяционным многочленом Эрмита) от этой матрицы: $g(Z) == H_g(Z)$, или, в подробной записи,

$$g(Z) = a_0 E + a_1 Z + \cdots + a_{n-1} Z^{n-1}, \qquad (1)$$

где E – единичная матрица, а коэффициенты $a_0, \dots, a_{n-1}$зависят как от функции g, так и от матрицы Z.

При этом верно следующее. Пусть функция $g(\lambda)$ комплексного переменного λ определена в области, содержащей все собственные числа $\lambda_1, \dots, \lambda_n$ матрицы Z и представляется там сходящимся степенным рядом

$$g(\lambda) = \gamma_0 + \gamma_1 \lambda + \gamma_2 \lambda^2 + \cdots. \qquad (2)$$

Тогда значение функции g от матрицы Z может быть получено подстановкой Z вместо λ в этот ряд:

$$g(Z) = \gamma_0 E + \gamma_1 Z + \gamma_2 Z^2 + \cdots, \qquad (3)$$

откуда следует, что элементы матрицы $g(Z)$ непрерывно зависят от элементов матрицы Z.

Нахождение $g(Z)$ возможно как вычисление некоторой частичной суммы ряда (3). Добавление каждого слагаемого этого ряда усложняет процесс вычисления на одно матричное умножение и одно матричное сложение (примерно n^3 обычных умножений и n^3 сложений). Ниже будет предложен алгоритм вычисления $g(Z)$, в котором учёт каждого следующего коэффициента γ_i усложняет процесс на $2n$ умножений и $2n$ сложений.

Прежде всего, если у матрицы Z все собственные числа различны, то интерполяционный многочлен Эрмита сводится к более простому многочлену Лагранжа

$A(\lambda) = a_0 + a_1 \lambda + \cdots + a_{n-1} \lambda^{n-1}$,

коэффициенты которого могут быть выражены явными формулами. Для простой записи этих формул условимся о следующем обозначении. Пусть $g_1(\lambda), \dots, g_n(\lambda)$ – функции комплексного аргумента λ, определённые во всех точках $\lambda_1, \dots, \lambda_n$. Тогда определитель порядка n, i-й столбец которого состоит из чисел $g_i(\lambda_1), \dots, g_i(\lambda_n)$, будем обозначать символом $[g_1, \dots, g_n]$. Так как при всех $p = 1, \dots, n$ имеем

$$a_0 + a_1 \lambda_p + \cdots + a_{n-1} \lambda_p^{n-1} = g(\lambda_p), \qquad (4)$$

то формулы для коэффициентов a_i таковы:

$$a_i = \frac{[1, \lambda, \dots, \lambda^{i-1}, g, \lambda^{i+1}, \dots, \lambda^{n-1}]}{[1, \lambda, \dots, \lambda^{n-1}]}. \qquad (5)$$

Знаменатель этих формул является определителем Вандермонда, который мы будем обозначать символом $W = W(\lambda_1, \dots, \lambda_n)$.

В статье будут рассматриваться только целые функции $g(\lambda)$, то есть такие, что ряд (2) сходится на всей комплексной плоскости. В частности, целыми являются матричные экспонента, синус и косинус, применяемые для решения систем линейных дифференциальных уравнений первого и второго порядка с постоянными коэффициентами.

Обозначим пространство размерности n точек $\Lambda = (\lambda_1, \dots, \lambda_n)$ с комплексными координатами через C^n, а через Ω^n – подмножество этого пространства, состоящее из точек с попарно различными координатами. Тогда, если $\Lambda \in \Omega^n$, то $W(\lambda_1, \dots, \lambda_n) \neq 0$, и формулы (5) однозначно определяют коэффициенты a_i. Однако, если функция $g(\lambda)$ целая, то, как мы увидим ниже, эти коэффициенты могут быть по непрерывности продолжены с множества Ω^n на всё пространство C^n. При этом все коэффициенты a_i окажутся симметрическими функциями переменных $\lambda_1, \dots, \lambda_n$, и, более того, целыми функциями от степенных сумм (сумм Ньютона) $s_m = \lambda_1^m + \cdots + \lambda_n^m$ при $m = 1, \dots, n$.

ЕВРАЗИЙСКОЕ НАУЧНОЕ ОБЪЕДИНЕНИЕ EURASIAN SCIENTIFIC ASSOCIATION
www.esa-conference.ru

Таким образом, если собственные числа матрицы Z некратны, то при всех $i = 0,1,\dots,n-1$ справедливы соотношения

$$a_i = a_i(g, s_1, \dots, s_n), \qquad (6)$$

в правых частях которых при фиксированной функции g стоят целые функции от n переменных. Так как множество матриц с некратными собственными числами всюду плотно в пространстве всех матриц порядка n, то формулы (6) могут быть продолжены по непрерывности на всё это пространство. Кроме того, как мы упоминали, матрица $g(Z)$ непрерывно зависит от матрицы Z. Поэтому матричнозначную функцию $g(Z)$ можно вычислять по формуле (1) с учётом (6) на всём пространстве квадратных комплексных матриц.

Суть предлагаемого способа вычисления целой функции от матрицы состоит в следующем. Сначала нужно с приемлемой точностью найти коэффициенты $a_0, \dots, a_{n-1}$, а затем вычислить $g(Z)$ по формуле (1). Переходим к обоснованию этого способа. Мы должны доказать сформулированное выше утверждение о возможности продолжения формул (5) с множества Ω^n на пространство C^n с получением целых функций от степенных сумм $s_1, \dots, s_n$. Но прежде мы докажем некоторое вспомогательное утверждение

Лемма. *Пусть $g(\lambda)$ – целая функция, $G(\lambda)$ – её первообразная, и пусть*

$$\Gamma(\Lambda) = G(\lambda_1) + \dots + G(\lambda_n) = \gamma_0 s_1 + \gamma_1 s_2/2 + \gamma_2 s_3/3 + \cdots. \qquad (7)$$

Тогда функция Γ является целой функцией от сумм $s_1, \dots, s_n$.

Доказательство. Кроме сумм Ньютона нам понадобятся элементарные симметрические многочлены $\sigma_1, \dots, \sigma_n$ от переменных $\lambda_1, \dots, \lambda_n$, связанные с суммами $s_1, \dots, s_n$ формулами Ньютона [2, с. 29], которые позволяют выразить $\sigma_1, \dots, \sigma_n$ как многочлены от сумм $s_1, \dots, s_n$ и наоборот. Вообще любой симметрический многочлен от $\lambda_1, \dots, \lambda_n$ является многочленом от $\sigma_1, \dots, \sigma_n$ или же от $s_1, \dots, s_n$. В частности, дальнейшие (при $m > n$) суммы s_m могут быть выражены через $s_1, \dots, s_n$ из соотношения

$$s_m - s_{m-1}\sigma_1 + \dots + (-1)^n s_{m-n}\sigma_n = 0, \qquad (8)$$

вследствие чего каждая частичная сумма ряда (7) является многочленом от $s_1, \dots, s_n$, и, следовательно, голоморфна по этим переменным. Для установления голоморфности общей суммы ряда по этим же переменным достаточно доказать, что он сходится, и притом равномерно, на любом кубе K_r: $|s_i| \le r$ при $i = 1, \dots, n$. Действительно, для любого наперёд заданного r найдётся такое $R > 0$, что на рассматриваемом кубе $|\sigma_i| < R^i, |s_i| < 2^i R^i$ для тех же i. Тогда из (8) по индукции легко вывести неравенство $|s_m| < 2^m R^m$ для всех $m > n$. Делается это так:

$$|s_m| < |s_{m-1}||\sigma_1| + \dots + |s_{m-n}||\sigma_n| < 2^m R^m.$$

Кроме того, так как функция $g(\lambda)$ целая, то в силу формулы Коши-Адамара

$$\lim_{m\to\infty} \sqrt[m]{|\gamma_m|} = 0,$$

что в сочетании с неравенством $|s_m| < 2^m R^m$ даёт

$$\lim_{m\to\infty} \sqrt[m]{|\gamma_m s_{m+1}|/(m+1)} = 0.$$

Поэтому ряд (7) равномерно сходится на любом кубе K_r, вследствие чего его сумма голоморфна по переменным $s_1, \dots, s_n$ на всём пространстве C^n, и лемма доказана.

Наконец, перейдём к доказательству продолжаемости по непрерывности коэффициентов a_i с множества Ω^n на пространство C^n и их голоморфности по переменным $s_1, \dots, s_n$. Мы докажем даже более общее утверждение, используя для простоты формул выражения $\tilde{s}_m = s_m/m$, которые назовём модифицированными суммами Ньютона. Разумеется, любой симметрический многочлен от $\lambda_1, \dots, \lambda_n$ является также многочленом от $\tilde{s}_1, \dots, \tilde{s}_n$. Вследствие доказанной леммы функция $\Gamma(\Lambda)$ является целой функцией от $\tilde{s}_1, \dots, \tilde{s}_n$.

Теорема. *Пусть $g_1(\lambda), \dots, g_n(\lambda)$ – целые функции комплексного переменного λ. Тогда частное от деления определителя $[g_1, \dots, g_n]$ на определитель Вандермонда W может быть по непрерывности продолжено с множества Ω^n на пространство C^n и является там целой функцией переменных $\tilde{s}_1, \dots, \tilde{s}_n$.*

Доказательство. Пусть $G_k(\lambda)$ – первообразная функция по отношению к $g_k(\lambda)$ при $k = 1, \dots, n$, и пусть $\Gamma_k(\lambda_1, \dots, \lambda_n) = G_k(\lambda_1) + \dots + G_k(\lambda_n)$. По доказанной лемме все Γ_k являются целыми функциями от модифицированных сумм $\tilde{s}_1, \dots, \tilde{s}_n$. Поэтому при любом $j = 1, \dots, n$ имеем

$$g_k(\lambda_j) = \frac{dG_k}{d\lambda_j} = \frac{\partial \Gamma_k}{\partial \lambda_j} = \sum_{q=1}^{n} \frac{\partial \Gamma_k}{\partial \tilde{s}_q}\frac{\partial \tilde{s}_q}{\partial \lambda_j} = \sum_{q=1}^{n} \lambda_j^{q-1}\frac{\partial \Gamma_k}{\partial \tilde{s}_q},$$

то есть элементы определителя $[g_1, \dots, g_n]$ являются суммами произведений элементов определителя Вандермонда на частные производные $\partial\Gamma_k/\partial\tilde{s}_q$. Отсюда следует разложение определителя $[g_1, \dots, g_n]$ в произведение определителей

$$[g_1, \dots, g_n] = W\begin{vmatrix} \partial\Gamma_1/\partial\tilde{s}_1 & \cdots & \partial\Gamma_n/\partial\tilde{s}_1 \\ \vdots & \ddots & \vdots \\ \partial\Gamma_1/\partial\tilde{s}_n & \cdots & \partial\Gamma_n/\partial\tilde{s}_n \end{vmatrix}.$$

Это доказывает теорему — ведь второй множитель в правой части полученного равенства является целой функцией от $\tilde{s}_1, \dots, \tilde{s}_n$.

Применительно к вычислению коэффициентов a_i мы должны взять $g_1(\lambda) = 1, g_2(\lambda) = \lambda, \dots, g_i(\lambda) = \lambda^{i-1}, g_{i+1}(\lambda) = g(\lambda), g_{i+2}(\lambda) = \lambda^{i+1}, \dots,$ $g_n(\lambda) = \lambda^{n-1},$ и несложное вычисление показывает, что $a_i = \partial\Gamma/\partial\tilde{s}_{i+1}$. А так как $\Gamma(\lambda_1, \dots, \lambda_n) = \gamma_0\tilde{s}_1 + \gamma_1\tilde{s}_2 + \cdots + \gamma_{n-1}\tilde{s}_n + \cdots$, то

$$a_i = \gamma_i + \sum_{m=n}^{\infty} \gamma_m \,\partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1}. \qquad (9)$$

Полученная формула уже достаточно проста и позволила бы последовательно накапливать с любой точностью коэффициенты $a_0, \dots, a_{n-1}$, если бы мы знали выражения для частных производных $\partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1}$ как многочленов от, например, степенных сумм $s_1, \dots, s_n$. Сами же степенные суммы можно было бы найти методом Леверье, то есть как следы степеней матрицы Z. Однако получающиеся при этом рекуррентные соотношения будут содержать коэффициенты характеристического многочлена матрицы Z, которые тоже ещё нужно вычислить. Поэтому здесь целесообразно использовать видоизменение метода Леверье, предложенное Д. К. Фаддеевым. Это позволит получить и явные формулы для производных $\partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1}$.

Напомним с точностью до обозначений суть метода Фаддеева [3, с. 295] для вычисления коэффициентов характеристического многочлена $\Delta_Z(\lambda)$ матрицы Z. Обозначим

$\Delta_Z(\lambda) = p_0\lambda^n + p_1\lambda^{n-1} + \cdots + p_n$,

где $p_0 = 1, p_1 = -\sigma_1, p_2 = \sigma_2, \dots, p_n = (-1)^n\sigma_n$. Пусть

$f_0(\lambda) = 1,\ f_1(\lambda) = \lambda + p_1,\ f_2(\lambda) = \lambda^2 + p_1\lambda + p_2, \dots,$

$f_{n-1}(\lambda) = \lambda^{n-1} + p_1\lambda^{n-2} + \cdots + p_{n-1},\ f_n(\lambda) = \Delta_Z(\lambda)$

– многочлены степеней $0, 1, \dots, n$ соответственно, которые мы будем называть многочленами Фаддеева. Вычисление коэффициентов $p_1, \dots, p_n$ идёт по следующей схеме:

$$F_i = f_i(Z), \qquad Z_{i+1} = ZF_i, \qquad p_{i+1} = -\frac{1}{i+1} trZ_{i+1} \qquad (10)$$

при $i = 0, 1, \dots, n-1$, где символом tr обозначается след матрицы. Многочлены $f_0, \dots, f_{n-1}$ линейно независимы, вследствие чего они образуют базис в пространстве всех многочленов степени не более $n-1$ (базис Фаддеева). Разложим по этому базису многочлен Лагранжа $A(\lambda)$:

$$a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} = b_0 + b_1 f_1(\lambda) + \cdots b_{n-1}f_{n-1}(\lambda). \qquad (11)$$

Подставляя сюда значения $\lambda = \lambda_1, \dots, \lambda_n$, мы вследствие формулы (4) можем найти выражения для коэффициентов $b_0, \dots, b_{n-1}$:

$$b_i = \frac{[1, f_1, \dots, f_{i-1}, g, f_{i+1}, \dots, f_{n-1}]}{[1, f_1, \dots, f_{n-1}]}, \quad i = 0, 1, \dots, n-1.$$

Выполняя как в числителе, так и в знаменателе преобразования второго рода для столбцов определителей (прибавление к одному столбцу другого столбца, умноженного на число), приходим к более простой формуле $b_i = [1, \lambda, \dots, \lambda^{i-1}, g, f_{i+1}, \dots, f_{n-1}]/W$. Отсюда видно, что каждый коэффициент b_i является линейной комбинацией коэффициентов $a_0, \dots, a_{n-1}$ с коэффициентами, являющимися симметрическими многочленами от $\lambda_1, \dots, \lambda_n$. Поэтому все коэффициенты $b_0, \dots, b_{n-1}$ могут быть продолжены по непрерывности с множества Ω^n на пространство C^n и являются целыми функциями от $s_1, \dots, s_n$.

Так как все коэффициенты b_i линейны по функции g, то, обозначив

$$b_i^{(m)} = [1, \lambda, \dots, \lambda^{i-1}, \lambda^m, f_{i+1}, \dots, f_{n-1}]/W, \qquad (12)$$

мы получим разложение коэффициента b_i в ряд

$$b_i = \gamma_0 b_i^{(0)} + \gamma_1 b_i^{(1)} + \gamma_2 b_i^{(2)} + \cdots, \qquad (13)$$

так что для вычисления b_i достаточно найти все компоненты $b_i^{(m)}$. Из (12) легко получаем

$$b_i^{(k)} = 0,\ \ k < i;\quad b_i^{(i)} = 1. \qquad (14)$$

Далее, рассматривая определитель $[1, \lambda, \dots, \lambda^{i-1}, f_m, f_{i+1}, \dots, f_{n-1}]$, равный, очевидно, нулю при $m = i+1, \dots, n-1$, деля его на W и учитывая (14), приходим к соотношениям $p_0 b_i^{(m)} + p_1 b_i^{(m-1)} + \dots + p_{m-i} b_i^{(i)} = 0$ при $i = 0,1,\dots,n-2$. Если условиться, что $b_i^{(k)} = 0$ при $k < 0$, то последнее соотношение можно записать в форме

$$p_0 b_i^{(m)} + p_1 b_i^{(m-1)} + \dots + p_n b_i^{(m-n)} = 0 \qquad (15)$$

при $i = 0,1,\dots,n-2$ и $m = i+1,\dots,n-1$.

Наконец, так как все числа $\lambda_1, \dots, \lambda_n$ являются корнями многочлена $Q_m(\lambda) = p_0\lambda^m + p_1\lambda^{m-1} + \dots + p_n\lambda^{m-n}$ при $m \geq n$, то определитель $[1, \lambda, \dots, \lambda^{i-1}, Q_m, f_{i+1}, \dots, f_{n-1}]$ равен нулю, откуда следует, что соотношение (15) верно при $i = 0,1,\dots,n-1$ и всех $m \geq i+1$. Итак, начальные значения компонент $b_i^{(m)}$ могут быть найдены согласно с равенствами (14), а для вычисления последующих значений следует применять формулы (15).

Мы уже упоминали два семейства симметрических многочленов − степенные суммы $s_1, s_2, \dots$ и элементарные симметрические многочлены $\sigma_0 = 1, \sigma_1, \dots, \sigma_n$. Теперь нам понадобится (бесконечное) семейство полных симметрических многочленов $h_0, h_1, h_2, \dots$. Многочлен h_m − это сумма всех возможных мономов степени m с коэффициентами, равными единице. Мы будем также считать, что $h_k = 0$ при $k < 0$. Тогда [2, с. 27] справедливо рекуррентное соотношение $h_m + p_1 h_{m-1} + p_2 h_{m-2} + \dots + p_n h_{m-n} = 0$ с начальным значением $h_0 = 1$. Сравнивая это с (15), получаем $b_i^{(m)} = h_{m-i}$. Отсюда и из (13) следует формула для вычисления коэффициентов b_i при $i = 0,1,\dots,n-1$:

$b_i = \gamma_i h_0 + \gamma_{i+1} h_1 + \gamma_{i+2} h_2 + \cdots$.

Таким образом, предлагаемый алгоритм вычисления целой функции $g(Z)$ состоит в следующем.

1. Выполняется алгоритм Фаддеева (10), в результате чего оказываются вычисленными матрицы $F_0, F_1, \dots, F_{n-1}$, а также коэффициенты $p_1, \dots, p_n$ характеристического многочлена $\Delta_Z(\lambda)$.
2. Определяются начальные значения $h_k = 0$ при $k < 0$, $h_0 = 1$; $b_i = \gamma_i$ при $i = 0,1,\dots,n-1$ и выполняются итерации

$h_m = -p_1 h_{m-1} - p_2 h_{m-2} - \dots - p_n h_{m-n}, \quad b_i := b_i + \gamma_{i+m} h_m$

при $m = 1,2,\dots$ до достижения достаточной точности в определении коэффициентов b_i. Каждая итерация требует $2n$ умножений и $2n$ сложений.

3. Функция $g(Z)$ вычисляется по формуле

$g(Z) = b_0 F_0 + b_1 F_1 + \dots + b_{n-1} F_{n-1}$.

В заключение дадим вывод формулы для частной производной $\partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1}$. Обозначим при $i = 0,1,\dots,n-1$ и всех $m \geq 0$ $a_i^{(m)} = [1, \lambda, \dots, \lambda^{i-1}, \lambda^m, \lambda^{i+1}, \dots, \lambda^{n-1}]/W$. Тогда вследствие (9) получаем при $i < n$, $m \geq 0$

$$a_i^{(m)} = \partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1}. \qquad (16)$$

Далее, если взять $g(\lambda) = \lambda^m$, то из (11) следует соотношение

$a_0^{(m)} + a_1^{(m)}\lambda + \dots + a_{n-1}^{(m)}\lambda^{n-1} = h_m + h_{m-1} f_1(\lambda) + \dots + h_{m-n+1} f_{n-1}(\lambda)$.

Приравнивая в этом тождестве коэффициенты при одинаковых степенях λ и учитывая (16), получаем

$\partial\tilde{s}_{m+1}/\partial\tilde{s}_{i+1} = p_0 h_{m-i} + p_1 h_{m-i-1} + \dots + p_{n-i-1} h_{m-n+1}$

для $m \geq n, i = 0,1,\dots,n-1$.

Литература:

1.Гантмахер Ф. Р. Теория матриц. М.: Наука, 1988.-552 с.
2.Макдональд И. Симметрические функции и многочлены Холла. М.: Мир, 1985.-222 с.
3.Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. – 2 изд. – М. – Л., 1963.-656 с.