ФУНКЦИИ ОТ МАТРИЦ. МАТРИЧНАЯ ЭКСПОНЕНТА

Рассмотрим квадратную матрицу A, $\dim A = n \times n$.

Определение 1. *Скалярной функцией* (СФМ) от квадратной матрицы A называется функция $f(A)$, которая реализует отображение

$$f(A)\colon R^{n\times n} \Longrightarrow R,$$

где R – множество действительных чисел.

Примеры:

детерминант, след, норма и число обусловленности матрицы.

Определение 2. *Векторной функцией* от квадратной матрицы A называется функция $f(A)$, которая реализует отображение

$$f(A): R^{n \times n} \Longrightarrow R^n,$$

где $R^n - n -$ мерное действительное пространство.

Примеры: векторы, построенные на элементах алгебраических спектров собственных значений и сингулярных чисел.

Матричные ряды и матричные функции от матриц

Матричная функция от матрицы (МФМ) реализует отображение

$$f(A): R^{n\times n} \Longrightarrow R^{n\times n}.$$

Определение 3. Пусть $f(\alpha)$ - скалярный степенной ряд (многочлен) относительно скалярной переменной α

$$f(A) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_p\alpha^p. \tag{1}$$

Тогда скалярный ряд $f(\alpha)$ порождает матричную функцию $f(A)$ от матрицы A в виде матричного ряда, если в представлении (1) для $f(\alpha)$ скалярную переменную заменить на матрицу A

$$f(A) = a_0I + a_1A + a_2A^2 + \cdots + a_pA^p. \tag{2}$$

Как осуществить переход от исходного представления МФМ в форме (2) к ее *минимальному представлению*, то есть к представлению матричным многочленом минимальной степени?

Теорема Гамильтона-Кэли. Квадратная матрица A с характеристическим полиномом

$$D(\lambda) = \det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

обнуляет свой характеристический полином так, что выполняется матричное соотношение

$$D(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI = 0 \tag{3}$$

где 0 - $(n \times n)$ нулевая матрица.

Пример. Рассмотрим матрицу

$$A = \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix}$$

Записываем характеристический полином $D(\lambda) = \det(\lambda I - A)$: $D(\lambda) = \lambda^3 + 6\lambda^2 + 11\lambda + 6$

Заменим λ на A в $D(\lambda)$:

$$D(A) = A^3 + 6A^2 + 11A + 6I =$$

$$= \begin{bmatrix} 70 & -116 & 19 \\ 71 & -117 & 19 \\ 64 & -102 & 11 \end{bmatrix} + \begin{bmatrix} -120 & 204 & -30 \\ -128 & 210 & -30 \\ -108 & 168 & -6 \end{bmatrix} + \begin{bmatrix} 44 & -88 & 11 \\ 55 & -99 & 11 \\ 44 & -66 & -11 \end{bmatrix} + \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

С помощью теоремы Гамильтона-Кэли введем следующие определения:

Определение 4. Многочлен (степенной ряд) $\varphi(\alpha)$ относительно скалярной переменной α называется *аннулирующим многочленом* квадратной матрицы A, если выполняется условие

$$\varphi(A) = 0. \tag{4}$$

Очевидно, *аннулирующим многочленом* матрицы A в силу теоремы Гамильтона-Кэли является в первую очередь ее *характеристический полином*.

Существует множество аннулирующих многочленов матрицы A степени большей, чем n. Но могут существовать аннулирующие многочлены степени $m < n$.

Определение 5 . Аннулирующий многочлен $\psi(\alpha)$ наименьшей степени m со старшим коэффициентом при α^m, равным единице, называется *минимальным многочленом* матрицы A.

Построим разложение многочлена $f(\alpha)$ (1), задающего матричную функцию от матрицы $f(A)$ в форме (2), по модулю минимального многочлена $\psi(\alpha)$ матрицы A, представив его выражением

$$f(\alpha) = \varphi(\alpha)\psi(\alpha) + r(\alpha). \quad (5)$$

где многочлен $r(\alpha)$ имеет степень $\deg(r(\alpha))$ меньше степени $\deg(\psi(\alpha))$ минимального многочлена $\psi(\alpha)$ матрицы A.

С помощью выражения (5) дадим следующее определение матричной функции от матрицы:

Определение 6. Пусть многочлен $f(\alpha)$ относительно скалярной переменной α допускает представление в форме (5), тогда матричная функция $f(A)$ может быть задана в *минимальной* форме

$$f(A) = r(A). \tag{6}$$

Проблема при задании матричной функции от матрицы в форме (6) – вычисление многочлена $r(\alpha)$.

Основной способ вычисления $r(\alpha)$ в силу (5) опирается на то, что $r(\alpha)$ является остатком от деления $f(\alpha)$ на минимальный многочлен

$$r(\alpha) = rest\,\frac{f(\alpha)}{\psi(\alpha)}. \tag{7}$$

Пример. Рассмотрим матрицу $A = \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix}$
с характеристическим полиномом $D(\lambda) = \lambda^3 + 6\lambda^2 + 11\lambda + 6$.

Требуется вычислить следующую матричную функцию

$$f(A) = A^4 + 4A^3 + 2A^2 - 10A - 10I$$

Заменим A на λ: $f(\lambda) = \lambda^4 + 4\lambda^3 + 2\lambda^2 - 10\lambda - 10$

Получим $f(\lambda) = \varphi(\lambda)\psi(\lambda) + r(\lambda)$ в соответствии с (5), где $\psi(\lambda) = D(\lambda)$.

Мы знаем, что $f(A) = \varphi(A)D(A) + r(A)$, где $D(A) = 0 \;\Rightarrow\; f(A) = r(A)$.

Вычислим $r(\lambda)$ как остаток от деления $\frac{f(\lambda)}{D(\lambda)}$. Получим $r(\lambda) = 3\lambda^2 + 6\lambda + 2$

Теперь $f(A) = A^4 + 4A^3 + 2A^2 - 10A - 10I$ может быть записана как $f(A) = 3A^2 + 6A + 2I$

Проверим наши вычисления в Matlab:

```
>> A=[4 -8 1; 5 -9 1; 4 -6 -1]

A =

     4    -8     1
     5    -9     1
     4    -6    -1

>> f_A=A^4+4*A^3+2*A^2-10*A-10*[1 0 0; 0 1 0; 0 0 1]

f_A =

   -34    54    -9
   -33    53    -9
   -30    48    -7

>> f_A1=3*A^2+6*A+2*[1 0 0; 0 1 0; 0 0 1]

f_A1 =

   -34    54    -9
   -33    53    -9
   -30    48    -7
```

Что делать, если $f(\alpha)$ *не является рядом или многочленом вида* (1), а является произвольной аналитической функцией со значениями на алгебраическом спектре собственных значений матрицы A?

В этом случае формирование матричной функции $f(A)$ от матрицы A опирается на представление $f(\alpha)$ в соответствии с интерполяционной схемой Лагранжа или в соответствии с интерполяционной схемой Ньютона.

Свойства матричной функции от матрицы

1. Матричная функция от матрицы $f(A)$ сохраняет спектр собственных векторов матрицы A: $A\xi_i = \lambda_i \xi_i$

$$f(A)\xi_i = f(\lambda_i)\xi_i \tag{8}$$

Пример

Пусть $f(\alpha) = \alpha^2,\ A = \begin{bmatrix} 4 & -8 & 1 \\ 5 & -9 & 1 \\ 4 & -6 & -1 \end{bmatrix}$

```
A=[4 -8 1;5  -9 1; 4 -6 -1]
[V,D]=eig(A)
A2=A^2
f_A_by_ksi=A2*V(:,1)
f_lambda_by_ksi=D(1,1)^2*V(:,1)
```

```
A =

     4    -8     1
     5    -9     1
     4    -6    -1


V =

    0.5774    0.8018    0.4082
    0.5774    0.5345    0.4082
    0.5774    0.2673    0.8165


D =

   -3.0000         0         0
         0   -1.0000         0
         0         0   -2.0000
```

```
A2 =

   -20    34    -5
   -21    35    -5
   -18    28    -1


f_A_by_ksi =

    5.1962
    5.1962
    5.1962


f_lambda_by_ksi =

    5.1962
    5.1962
    5.1962
```

2. Матричная функция от матрицы $f(A)$ сохраняет матричное соотношение подобия, т.е. если $B = T^{-1}AT$, то

$$f(B) = T^{-1}f(A)T. \tag{9}$$

3. Матричная функция от матрицы $f(A)$ сохраняет блочно-диагональную форму матрицы A, т.е. если $A = diag\{a_i\}$, то

$$f(A) = diag\{f(a_i)\}. \tag{10}$$

Матричная экспонента

Скалярный ряд

$$e^{\alpha t} = 1 + \alpha t + \frac{(\alpha t)^2}{2!} + \frac{(\alpha t)^3}{3!} + \cdots + \frac{(\alpha t)^p}{p!} + \cdots = \sum_{k=0}^{\infty} \frac{(\alpha t)^k}{k!} \quad (11)$$

Заменяем α на A:

$$e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \cdots + \frac{(At)^p}{p!} + \cdots = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} \quad (12)$$

Основные способы вычисления матричной экспоненты:

1. Численный способ.

Основан на переходе от непрерывного времени t к дискретному k, выраженному числом интервалов дискретности длительности Δt $\quad t = k(\Delta t)$.

$$e^{At} = e^{A\Delta t k} = (e^{A\Delta t})^k = (\bar{A})^k$$

$$\bar{A} = I + A\Delta t + \frac{1}{2!}(A\Delta t)^2 + \frac{1}{3!}(A\Delta t)^3 + \cdots + \frac{1}{p!}(A\Delta t)^p$$

Пример.

$$A = \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix}$$

$$\Delta t = 0{,}1$$

$$k = 1$$

$$p = 2$$

$$e^{At} = e^{A\Delta t k} = (e^{A\Delta t})^{k} = (\bar{A})^{1} = \bar{A}$$

$$\bar{A} = I + A\Delta t + \frac{1}{2!}(A\Delta t)^{2} = \begin{bmatrix} 1{,}22 & -0{,}21 \\ 0 & 0{,}905 \end{bmatrix}$$

2. Метод диагонализации (или метод собственных значений)

Применяется к матрицам простой структуры, для которых справедливо соотношение $M\Lambda = AM$

$$e^{At} = Me^{\Lambda t}M^{-1} = Mdiag\{e^{\lambda_i t}, i = 1 \dots n\}M^{-1}, \qquad (12)$$

$$M = row\{M_i = \xi_i, i = 1, \dots n\}$$

M – матрица собственных векторов матрицы A.

Пример.

$$A = \begin{bmatrix} 2 & -2 \\ 0 & 1 \end{bmatrix}$$

Вычислим собственные значения: $\lambda_1 = -1, \lambda_2 = 2$.

Вычислим собственные векторы: $\xi_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \xi_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Составим матрицу $M = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}$ и $M^{-1} = \begin{bmatrix} 0 & \frac{1}{3} \\ 1 & -\frac{2}{3} \end{bmatrix}$.

Подставляем в формулу $e^{At} = Mdiag\{e^{\lambda_i t}, i = 1 \ldots n\}M^{-1}$

$$e^{At} = \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{3} \\ 1 & -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} e^{2t} & \frac{2}{3}e^{-t} - \frac{2}{3}e^{2t} \\ 0 & e^{-t} \end{bmatrix}$$

Если подставить $t = 0{,}1$, получим

$$e^{At} = \begin{bmatrix} e^{2t} & \frac{2}{3}e^{-t} - \frac{2}{3}e^{2t} \\ 0 & e^{-t} \end{bmatrix} = \begin{bmatrix} 1{,}2214 & -0{,}211 \\ 0 & 0{,}9048 \end{bmatrix}$$

3. Способ, основанный на приведении к нормальной Жордановой форме

Применяется к матрицам, спектр собственных значений которых содержит r кратных собственных значений λ_i кратности m_i каждое. Тогда выполняется матричное соотношение подобия

$$TJ = AT,$$

$$J = diag\left\{ J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}$$

В результате для матричной экспоненты можно записать

$$e^{At} = Te^{Jt}T^{-1} \tag{13}$$

$$e^{Jt} = diag\left\{ e^{J_i t} = \begin{bmatrix} e^{\lambda_i t} & \frac{te^{\lambda_i t}}{1!} & \frac{t^2 e^{\lambda_i t}}{2!} & \dots & \frac{t^{n-1} e^{\lambda_i t}}{(m_i - 1)!} \\ 0 & e^{\lambda_i t} & \frac{te^{\lambda_i t}}{1!} & \dots & \frac{t^{n-2} e^{\lambda_i t}}{(m_i - 2)!} \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & \frac{te^{\lambda_i t}}{1!} \\ 0 & 0 & 0 & \dots & e^{\lambda_i t} \end{bmatrix}; \sum_{i=1}^{r} m_i = n \right\}$$

Пример.

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 2 & 24 \\ 0 & 1 & -2 \end{bmatrix}$$

Собственные значения $\lambda_1 = \lambda_2 = \lambda_3 = -1$.

Матрицы $T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix}, T^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -2 \end{bmatrix}$

Матрица $J = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$

$$e^{At} = Te^{Jt}T^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-t} & te^{-t} & \frac{1}{2}t^2e^{-t} \\ 0 & e^{-t} & te^{-t} \\ 0 & 0 & e^{-t} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -2 \end{bmatrix}$$

4. Способ преобразования Лапласа

Вычисление обратного преобразования Лапласа от резолвенты $(sI - A)^{-1}$ в форме

$$e^{At} = \mathcal{L}^{-1}\{(sI - A)^{-1}\} \tag{14}$$

Для разложения резольвенты без ее обращения используется алгоритм Фаддеева-Леверье на основе представления

$$(sI - A)^{-1} = \frac{1}{\det(sI - A)}\left[\Delta(sI - A)\right]^T = \frac{s^{n-1}H_0 + s^{n-2}H_1 + \ldots + H_{n-1}}{s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n} \tag{15}$$

$(n \times n)$ − матрицы H_i $(i = 0, \dots n-1)$ и коэффициенты характеристического уравнения вычисляются с помощью рекуррентной процедуры

$$\begin{aligned} H_0 &= I, & a_1 &= -tr(AH_0) \\ H_1 &= AH_0 + a_1 I, & a_2 &= -tr(AH_1)/2 \\ &\dots\dots\dots\dots & &\dots\dots\dots\dots \\ H_k &= AH_{k-1} + a_k I, & a_{k+1} &= -tr(AH_k)/k \end{aligned} \tag{16}$$

Теперь резолвенту можно переписать как

$$(sI - A)^{-1} = \frac{s^{n-1}}{D(s)} H_0 + \frac{s^{n-1}}{D(s)} H_1 + \dots + \frac{s}{D(s)} H_{n-2} + \frac{1}{D(s)} H_{n-1} \tag{17}$$

И матричная экспонента принимает вид

$$e^{At} = L^{-1}\left\{\frac{s^{n-1}}{D(s)}\right\} H_0 + L^{-1}\left\{\frac{s^{n-1}}{D(s)}\right\} H_1 + \dots + L^{-1}\left\{\frac{s}{D(s)}\right\} H_{n-2} + L^{-1}\left\{\frac{1}{D(s)}\right\} H_{n-1} \tag{18}$$

Пример. Вычислить матричную экспоненту с помощью преобразования Лапласа

Алгоритм вычисления:

1) Вычислить матрицу $sI - A$
2) Вычислить H_k и a_k
3) Записать резолвенту в виде(17)
4) Вычислить обратное преобразование Лапласа от резолвенты

Рассмотрим $A = \begin{bmatrix} 0 & 1 \\ -10 & -7 \end{bmatrix}$

1) Матрица $sI - A = \begin{bmatrix} s & -1 \\ 10 & s+7 \end{bmatrix}$

2) Вычислим $a_1 = -tr(AH_0), H_0 = I \Rightarrow a_1 = -tr(A)$

Характеристический полином A $D(s) = s^2 + 7s + 10$

Его корни: $\lambda_1 = -2, \lambda_2 = -5 \Rightarrow a_1 = -tr(A) = 7,$

$a_2 = -tr(AH_1)/2, H_1 = AH_0 + a_1 I$

$$H_1 = \begin{bmatrix} 0 & 1 \\ -10 & -7 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} 7 & 1 \\ -10 & 0 \end{bmatrix}$$

$$AH_1 = \begin{bmatrix} 0 & 1 \\ -10 & -7 \end{bmatrix} \begin{bmatrix} 7 & 1 \\ -10 & 0 \end{bmatrix} = \begin{bmatrix} -10 & 0 \\ -140 & -10 \end{bmatrix}$$

$$a_2 = -\frac{tr(AH_1)}{2} = -\frac{-20}{2} = 10$$

3) $(sI - A)^{-1} = \frac{1}{D(s)}[sH_0 + H_1] = \frac{1}{s^2+7s+10}\left[s\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 7 & 1 \\ -10 & 0 \end{bmatrix}\right] =$

$$= \frac{1}{s^2 + 7s + 10}\left[\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} + \begin{bmatrix} 7 & 1 \\ -10 & 0 \end{bmatrix}\right] = \begin{bmatrix} \dfrac{s+7}{(s+2)(s+5)} & \dfrac{1}{(s+2)(s+5)} \\ \dfrac{-10}{(s+2)(s+5)} & \dfrac{s}{(s+2)(s+5)} \end{bmatrix}$$

4) Вычисляем обратное преобразование Лапласа:

$$\mathcal{L}^{-1}\left\{\frac{s+7}{(s+2)(s+5)}\right\} = \sum_{i=1}^{2} \frac{s_i+7}{2s_i+7} e^{s_i} = \frac{5}{3}e^{-2t} - \frac{2}{3}e^{-5t}$$

$$\mathcal{L}^{-1}\left\{\frac{-10}{(s+2)(s+5)}\right\} = -10\sum_{i=1}^{2} \frac{1}{2s_i+7} e^{s_i} = -\frac{10}{3}e^{-2t} + \frac{10}{3}e^{-5t}$$

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+2)(s+5)}\right\} = \sum_{i=1}^{2} \frac{1}{2s_i+7} e^{s_i} = -\frac{1}{3}e^{-2t} + \frac{1}{3}e^{-5t}$$

$$\mathcal{L}^{-1}\left\{\frac{s}{(s+2)(s+5)}\right\} = \sum_{i=1}^{2} \frac{s_i}{2s_i+7} e^{s_i} = -\frac{2}{3}e^{-2t} + \frac{5}{3}e^{-5t}$$

Окончательно получаем

$$e^{At} = \begin{bmatrix} \frac{5}{3}e^{-2t} - \frac{2}{3}e^{-5t} & -\frac{1}{3}e^{-2t} + \frac{1}{3}e^{-5t} \\ -\frac{10}{3}e^{-2t} + \frac{10}{3}e^{-5t} & -\frac{2}{3}e^{-2t} + \frac{5}{3}e^{-5t} \end{bmatrix}$$

Обращение матриц с помощью теоремы Гамильтона - Кэли

Запишем матричное соотношение (3), представляющее собой аналитическое содержание теоремы Гамильтона – Кэли, в форме

$$A^n + a_1 A^{n-1} + \cdots + a_{n-1} A + a_n I = 0 \tag{19}$$

Умножим (19) слева на A^{-1}:

$$a_n A^{-1} + a_{n-1} I + a_{n-2} A + \cdots + a_1 A^{n-2} + A^{n-1} = 0 \tag{20}$$

Разрешим (20) относительно обратной матрицы

$$A^{-1} = -(a_n)^{-1}(a_{n-1} I + a_{n-2} A + \cdots + a_1 A^{n-2} + A^{n-1}) =$$

$$= -(a_n)^{-1}\left(A^{n-1} + \sum_{i=1}^{n-1} a_i A^{n-1-i}\right) \tag{21}$$

Вывод.

Получили алгоритмическую базу обращения матриц.

Достоинство способа: обращение матриц с помощью приведенного матричного соотношения нечувствительно к обусловленности обращаемой матрицы.

Недостаток: необходимость знания коэффициентов характеристического полинома.

Предлагаемая процедура обращения не вызовет заметных сложностей для случая разреженных матриц, и особенно удобно пользоваться ею при обращении матриц заданных в сопровождающей (фробениусовой) форме, потому что коэффициенты характеристического полинома в явном виде присутствуют в ней.

Пример.

$$A = \begin{bmatrix} 0 & 0 & -10 \\ 1 & 0 & -7 \\ 0 & 1 & 4 \end{bmatrix}$$

Характеристический полином: $D(\lambda) = \lambda^3 - 4\lambda^2 + 7\lambda + 10,$

$$a_1 = -4,\, a_2 = 7, a_1 = 10$$

$$A^{-1} = -(a_3)^{-1}(A^2 + a_1 A + a_2 A^0)$$

После подстановки получаем: $A^{-1} = \begin{bmatrix} -7 & 1 & 0 \\ 4 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$