Тема 6.

Элементы теории матриц (продолжение)

В дальнейшем мы будем широко пользоваться операциями дифференцирования и интегрирования различных матричных функций скалярного аргумента по этому аргументу. Такие операции выполняются поэлементно:

$$\mathbf{F}'(t)=\begin{pmatrix} F'_{11}(t) & F'_{12}(t) & . & . & F'_{1m}(t) \\ F'_{21}(t) & F'_{22}(t) & . & . & F'_{2m}(t) \\ . & . & . & . & . \\ . & . & . & . & . \\ F'_{m1}(t) & F'_{m2}(t) & . & . & F'_{mm}(t) \end{pmatrix};$$

$$\int \mathbf{F}(t)dt=\begin{pmatrix} \int F_{11}(t)dt & \int F_{12}(t)dt & . & . & \int F_{1m}(t)dt \\ \int F_{21}(t)dt & \int F_{22}(t)dt & . & . & \int F_{2m}(t)dt \\ . & . & . & . & . \\ . & . & . & . & . \\ \int F_{m1}(t)dt & \int F_{m2}(t)dt & . & . & \int F_{mm}(t)dt \end{pmatrix}.$$

В частности справедливы следующие формулы:

$$\frac{d}{dt}\left(\mathbf{F}(t)\cdot\mathbf{G}(t)\right)=\frac{d\mathbf{F}(t)}{dt}\mathbf{G}(t)+\mathbf{F}(t)\frac{d\mathbf{G}(t)}{dt};$$

$$\frac{d\mathbf{F}^k(t)}{dt}=\frac{d\mathbf{F}(t)}{dt}\mathbf{F}^{k-1}(t)+\mathbf{F}(t)\frac{d\mathbf{F}(t)}{dt}\mathbf{F}^{k-2}(t)+\ldots+\mathbf{F}^{k-1}(t)\frac{d\mathbf{F}(t)}{dt};$$

$$\frac{d\mathbf{F}^{-1}(t)}{dt}= \ ???$$

$$\mathbf{F}(t)\cdot\mathbf{F}^{-1}(t)=\mathbf{E}$$

Продифференцируем это равенство:

$$\frac{d}{dt}\left(\mathbf{F}(t)\cdot\mathbf{F}^{-1}(t)\right)=\mathbf{0}; \quad \Rightarrow \quad \cdot\frac{d}{dt}\mathbf{F}(t)\cdot\mathbf{F}^{-1}(t)+\mathbf{F}(t)\cdot\frac{d}{dt}\mathbf{F}^{-1}(t)=\mathbf{0}$$

$$\frac{d\mathbf{F}^{-1}(t)}{dt} = -\mathbf{F}^{-1}(t)\frac{d\mathbf{F}(t)}{dt}\mathbf{F}^{-1}(t).$$

Нормы матриц

Для вводимой величины – нормы матрицы, будем использовать обозначение $\|\mathbf{A}\|$. Введем это понятие аксиоматически. В дальнейшем пусть α - скалярная величина. *Нормой матрицы* $\|\mathbf{A}\|$ называется число, удовлетворяющее следующим четырем аксиомам:

1. $\|\mathbf{A}\| \geq 0$, при этом $\|\mathbf{A}\| = 0$ тогда и только тогда, когда $\mathbf{A}=\mathbf{0}$;
2. $\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|$;
3. $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$;
4. $\|\mathbf{A}\cdot\mathbf{B}\| \leq \|\mathbf{A}\|\cdot\|\mathbf{B}\|$;

Если дополнительно выполняются еще две аксиомы, то такая норма называется *канонической*:

5. $|a_{ik}| \leq \|\mathbf{A}\|; \quad \forall i,k$
6. Если $\forall i,k \quad |a_{ik}| \leq |b_{ik}|$, то $\|\mathbf{A}\| \leq \|\mathbf{B}\|$.

Примерами матричной нормы являются:

$$\|\mathbf{A}\|_1 = \max_i \sum_{j=1}^{n} |a_{ij}|; \quad \|\mathbf{A}\|_2 = \max_j \sum_{i=1}^{n} |a_{ij}|; \quad \|\mathbf{A}\|_3 = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$ – евклидова норма.

Не трудно показать, что все эти нормы являются каноническими и удовлетворяют всем шести аксиомам. *Вопрос: «Можно между этими нормами для произвольной матрицы поставить какие-то знаки неравенства?»*. Правильный ответ на него иллюстрирует следующий пример.

Пример. Определить все три нормы для двух следующих матриц:

$$\mathbf{A} = \begin{pmatrix} -1 & 2 & -3 \\ 4 & -5 & 6 \\ -7 & 8 & -9 \end{pmatrix}; \quad \mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Непосредственные вычисления дают:

$$\|\mathbf{A}\|_1 = 24; \quad \|\mathbf{A}\|_2 = 18; \quad \|\mathbf{A}\|_3 \simeq 16.9; \quad \|\mathbf{E}\|_1 = 1; \quad \|\mathbf{E}\|_2 = 1; \quad \|\mathbf{E}\|_3 = \sqrt{3}.$$

Матричный ряд и матричные функции

Наиболее простыми матричными функциями являются полиномы. Для построения полинома необходимо образовать степени матрицы **A**: $\mathbf{A}^0$, $\mathbf{A}^1$, $\mathbf{A}^2$, ..., $\mathbf{A}^n$. Степени матриц определены лишь для квадратных матриц, и поэтому в дальнейшем речь будет идти только о них. По определению $\mathbf{A}^k = \mathbf{A\ A\ A} \ldots \mathbf{A}$ (матрица умножается k раз) и $\mathbf{A}^0=\mathbf{E}$. Тогда матричный полином со скалярными коэффициентами c_0, c_1, ..., c_n будет выглядеть так:

$$\mathbf{P}_n(\mathbf{A}) = c_0\mathbf{E} + c_1\mathbf{A} + c_2\mathbf{A}^2 + \ldots + c_n\mathbf{A}^n .$$

Аргументом (область определения) является квадратная матрица размерности $m \times m$, и значением будет матрица (область значений) той же размерности.

Теперь устремим n к бесконечности, т.е. формально перейдем к бесконечной сумме

$$\mathbf{P}(\mathbf{A}) = \sum_{\gamma=0}^{\infty} c_\gamma \mathbf{A}^\gamma . \tag{1}$$

Такая сумма называется степенным матричным рядом относительно матрицы **A**. Матричному ряду естественно сопоставить скалярный ряд

$$p(x) = \sum_{\gamma=0}^{\infty} c_\gamma x^\gamma .$$

Матричный ряд будем называть сходящимся, если сходятся все m^2 скалярных рядов для элементов матрицы $\mathbf{P}(\mathbf{A})$. Введенное нами понятие нормы позволяет установить достаточное условие сходимости матричного ряда. Введем матрицу $\mathbf{U}^{(\gamma)} = c_\gamma \mathbf{A}^\gamma$. Обозначим ее элементы за $u_{kj}^{(\gamma)}$, а элементы матрицы $\mathbf{P}(\mathbf{A})$ за $p_{k,j}$. Тогда с учетом выполнения шести аксиом для канонической нормы имеем цепочку неравенств

$$\left|p_{k,j}\right| = \left|\sum_{\gamma=0}^{\infty} u_{kj}^{(\gamma)}\right| \le \sum_{\gamma=0}^{\infty}\left|u_{kj}^{(\gamma)}\right| \le \sum_{\gamma=0}^{\infty}\left\|c_\gamma \mathbf{A}^\gamma\right\| = \sum_{\gamma=0}^{\infty}\left|c_\gamma\right|\left\|\mathbf{A}^\gamma\right\| \le \sum_{\gamma=0}^{\infty}\left|c_\gamma\right|\left\|\mathbf{A}\right\|^\gamma \tag{2}$$

В результате *достаточным* условием сходимости матричного ряда (1) является выполнения условия

$$\|\mathbf{A}\| < R, \tag{3}$$

являющегося, в свою очередь, условием абсолютной сходимости скалярного степенного ряда, стоящего последним в цепочке (2). Здесь R – радиус сходимости скалярного степенного ряда.

Правильному пониманию условия (3) способствует ответ на следующие два вопроса. Пусть имеется ряд с R=17. Сходится ли матричный ряд (1), если

а) $\|\mathbf{A}\|_1 = 24$, $\quad \|\mathbf{A}\|_2 = 18$, $\quad \|\mathbf{A}\|_3 = 16.9$;

б) $\|\mathbf{A}\|_1 = 24, \quad \|\mathbf{A}\|_2 = 18, \quad \|\mathbf{A}\|_3 = 17.5$.

Необходимое условие сходимости матричного ряда рассмотрим несколько позже.

Если матричный ряд сходится, то матрицу $\mathbf{P}(\mathbf{A})$ будем называть *матричной функцией* (пока ограничимся матричными функциями только такого вида). Примерами матричных функций могут служить

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}, \qquad \cos(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{(-1)^k \mathbf{A}^{2k}}{(2k)!},$$

$$\sin(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{(-1)^k \mathbf{A}^{2k+1}}{(2k+1)!}, \qquad (\mathbf{E} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k .$$

Первые три ряда (матричные экспонента, косинус и синус), так же, как и их скалярные аналоги, сходятся для любых матриц, а последний ряд (геометрическая прогрессия) имеет R=1. Попутно легко убедиться, что $\sin^2(\mathbf{A}) + \cos^2(\mathbf{A}) = \mathbf{E}$.

Для дальнейшего изложения необходимо ввести понятие подобных матриц.

Определение. Пусть задана матрица **A** и некоторая неособенная матрица **S** (т.е. $\det(\mathbf{S}) \neq 0$ и существует $\mathbf{S}^{-1}$). Всякая матрица $\mathbf{B}=\mathbf{SAS}^{-1}$ называется подобной матрице **A**. Очевидно, что и $\mathbf{A} = \mathbf{S}^{-1}\mathbf{BS}$ подобна **B**.

С точки зрения линейных преобразований можно сказать, что две матрицы подобны, если они соответствуют одному и тому же линейному преобразованию в различных базисах. Поясним, что это означает. В некотором базисе вектор **x** преобразуется в вектор ***y*** посредством матрицы **A***:* **y**=**Ax**. Переход к новому базису осуществляет матрица **S**, т.е. образы векторов **x** и **y** в новом базисе имеют вид: $\xi=\mathbf{Sx}$ и $\upsilon=\mathbf{Sy}$. Умножив оба равенства слева на $\mathbf{S}^{-1}$, получим $\mathbf{x}=\mathbf{S}^{-1}\xi$ и $\mathbf{y}=\mathbf{S}^{-1}\upsilon$. В новом базисе равенство **y**=**Ax** превращается в $\mathbf{S}^{-1}\upsilon=\mathbf{AS}^{-1}\xi$ или $\upsilon=\mathbf{SAS}^{-1}\xi$. Это и означает, что матрица $\mathbf{SAS}^{-1}$ (подобная матрице **A**) осуществляет то же самое линейное преобразование, что и **A**, но в другом базисе.

Сформулируем теперь несколько теорем о подобных матрицах и матричных функциях, устанавливающих некоторые свойства тех и других.

Теорема 1. Подобные матрицы **A** и $\mathbf{B} =\mathbf{SAS}^{-1}$ имеют одинаковые собственные значения. При этом, если собственному значению λ матрицы **A** отвечает собственный вектор **u**, то у матрицы **B** этому же собственному числу λ соответствует собственный вектор **Su**.

Доказательство. Так как $\mathbf{SS}^{-1} = \mathbf{E}$, то $\det(\mathbf{SS}^{-1}) = \det(\mathbf{S}) \det(\mathbf{S}^{-1}) = \det(\mathbf{E}) = 1$. Для характеристических полиномов **A** и **B** имеем

$$\det(\mathbf{B} - \lambda\mathbf{E}) = \det(\mathbf{SAS}^{-1} - \lambda\mathbf{SS}^{-1}) = \det(\mathbf{S}(\mathbf{A} - \lambda\mathbf{E})\mathbf{S}^{-1}) =$$

$$= \det(\mathbf{S}) \cdot \det\left(\mathbf{A} - \lambda \mathbf{E}\right) \cdot \det(\mathbf{S}^{-1}) = \det\left(\mathbf{A} - \lambda \mathbf{E}\right).$$

Характеристические полиномы для обеих матриц совпали, следовательно, совпали и их корни, т.е. собственные значения. Для доказательства второй части теоремы в равенстве $\mathbf{Au} = \lambda \mathbf{u}$ заменим матрицу $\mathbf{A}$ на подобную ей $\mathbf{S}^{-1}\mathbf{BS}$:

$$\mathbf{S}^{-1}\mathbf{BSu} = \lambda \mathbf{u}.$$

Теперь, умножив обе части равенства на $\mathbf{S}$ слева, получим требуемый результат:

$$\mathbf{B}\left(\mathbf{Su}\right) = \lambda\left(\mathbf{Su}\right).$$

<u>*Теорема* 2</u>. Если матрицы $\mathbf{A}$ и $\mathbf{B}$ подобны, то их матричные функции также подобны.

Иными словами, если $\mathbf{B}=\mathbf{SAS}^{-1}$, то $\mathbf{f}(\mathbf{B})=\mathbf{Sf}(\mathbf{A})\mathbf{S}^{-1}$.

<u>*Доказательство*</u>. Первоначально определим $\mathbf{B}^k$.

$$\mathbf{B}^k = \left(\mathbf{SAS}^{-1}\right)^k = \mathbf{SAS}^{-1}\mathbf{SAS}^{-1}\ldots\mathbf{SAS}^{-1} = \mathbf{SAA}\ldots\mathbf{AS}^{-1} = \mathbf{SA}^k\mathbf{S}^{-1},$$

$$\mathbf{f}(\mathbf{B}) = \sum_{k=0}^{\infty} c_k \mathbf{B}^k = \mathbf{S}\left(\sum_{k=0}^{\infty} c_k \mathbf{A}^k\right)\mathbf{S}^{-1} = \mathbf{Sf}(\mathbf{A})\mathbf{S}^{-1}.$$

<u>*Теорема* 3</u>. Матрица $\mathbf{A}$ с различными собственными значениями $\lambda_1, \lambda_2, \ldots, \lambda_m$ (нет кратных!) подобна некоторой диагональной матрице $\mathbf{\Lambda}$, на главной диагонали которой стоят собственные значения матрицы $\mathbf{A}$, т.е. $\mathbf{\Lambda} = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_m\right)$.

<u>*Доказательство*</u>. Пусть $\mathbf{u}_k$ – собственные векторы матрицы $\mathbf{A}$. Обозначим за $\mathbf{U}$ матрицу, столбцами которой являются все $\mathbf{u}_k$. Тогда

$$\mathbf{AU} = \mathbf{A}\begin{pmatrix} | & | & & & | \\ | & | & & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \ldots & \mathbf{u}_m \\ | & | & & & | \\ | & | & & & | \end{pmatrix} = \begin{pmatrix} | & | & & & | \\ | & | & & & | \\ \lambda_1\mathbf{u}_1 & \lambda_2\mathbf{u}_2 & \ldots & \ldots & \lambda_m\mathbf{u}_m \\ | & | & & & | \\ | & | & & & | \end{pmatrix} =$$

$$= \begin{pmatrix} | & | & & & | \\ | & | & & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \ldots & \mathbf{u}_m \\ | & | & & & | \\ | & | & & & | \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & & \lambda_m \end{pmatrix} = \mathbf{U\Lambda}.$$

Умножая полученное равенство поочередно справа и слева на $\mathbf{U}^{-1}$, получаем требуемое

$$\mathbf{\Lambda} = \mathbf{U}^{-1}\mathbf{AU}, \qquad\qquad \mathbf{A} = \mathbf{U\Lambda U}^{-1}.$$

Попутно заметим, что мы одновременно не только доказали теорему, но и определили матрицу преобразования подобия $\mathbf{U}$, состоящую из линейно независимых столбцов $\mathbf{u}_k$ и, следовательно, неособенную. Не вдаваясь в подробности, отметим, что в общем случае, при наличии кратных собственных значений у матрицы $\mathbf{A}$, уже не имеющей простую структуру, вместо матрицы $\mathbf{\Lambda}$, возникает клеточно-диагональная матрица, где каждая клетка представляет собой так называемый *канонический ящик Жордана*.

Исключительно для простоты изложения дальнейшие теоремы будут доказываться только для матриц простой структуры, однако результаты справедливы и для более общего случая.

Теорема 4. Если собственные значения матрицы $\mathbf{A}$ обозначить через λ_1, λ_2, ..., λ_m, то собственными значениями матрицы $\mathbf{f}(\mathbf{A})$ будут числа $f(\lambda_1)$, $f(\lambda_2)$, …, $f(\lambda_m)$.

Доказательство. $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$. По *теореме* 2 $\mathbf{f}(\mathbf{A}) = \mathbf{U}^{-1}\mathbf{f}(\mathbf{\Lambda})\mathbf{U}$. Представим $\mathbf{f}(\mathbf{\Lambda})$ в покомпонентном виде

$$\mathbf{f}(\mathbf{\Lambda}) = \sum_{k=0}^{\infty} c_k \mathbf{\Lambda}^k = \begin{pmatrix} \sum\limits_{k=0}^{\infty} c_k \lambda_1^k & 0 & . & . & 0 \\ 0 & \sum\limits_{k=0}^{\infty} c_k \lambda_2^k & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \sum\limits_{k=0}^{\infty} c_k \lambda_m^k \end{pmatrix} =$$

$$= \begin{pmatrix} f(\lambda_1) & 0 & . & . & 0 \\ 0 & f(\lambda_2) & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & f(\lambda_m) \end{pmatrix}$$

И тогда

$$\mathbf{f}(\mathbf{A}) = \mathbf{U}^{-1} f(\mathbf{\Lambda})\mathbf{U} = \mathbf{U}^{-1} \begin{pmatrix} f(\lambda_1) & 0 & . & . & 0 \\ 0 & f(\lambda_2) & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & f(\lambda_m) \end{pmatrix} \mathbf{U}.$$

что и означает, что $f(\lambda_1)$, $f(\lambda_2)$, …, $f(\lambda_m)$ являются собственными значениями матрицы $\mathbf{f}(\mathbf{A})$, так как преобразование подобия не меняет собственных значений.

Следствие 1. Из вышеприведенных формул непосредственно следует, что матричный ряд $\mathbf{f}(\mathbf{A})$ существует тогда и только тогда, когда существуют _все_ скалярные степенные ряды, стоящие на диагонали матрицы $\mathbf{f}(\mathbf{\Lambda})$, а у тех, в свою очередь, необходимым и достаточным условием существования является выполнение условий

$$|\lambda_k| < R, \qquad \forall \lambda_k \tag{4}$$

Таким образом, условие (4) является необходимым и достаточным условием сходимости матричного степенного ряда. На практике вопрос о сходимости матричного ряда решается в такой последовательности: 1) сначала находится радиус сходимости R соответствующего скалярного ряда, 2) а затем проверяется выполнение условия (4) для всех собственных значений.

Следствие 2. Поскольку условие (3) является лишь достаточным условием сходимости матричного степенного ряда, а условие (4) необходимым и достаточным, то из совместного рассмотрения обоих условий легко заключить, что

$$|\lambda_k| \le \|\mathbf{A}\| \,, \tag{5}$$

т.е., что все собственные значения матрицы не превышают ее любую каноническую норму. Формула (5) будет многократно использоваться нами в дальнейшем. В ее справедливости рекомендуется убедиться самостоятельно.

Теорема 5. Две любые функции матрицы $\mathbf{A}$ коммутируют между собой:

$$\mathbf{f}(\mathbf{A}) \cdot \mathbf{g}(\mathbf{A}) = \mathbf{g}(\mathbf{A}) \cdot \mathbf{f}(\mathbf{A}).$$

Доказательство. По теореме 2

$$\mathbf{f}(\mathbf{A}) = \mathbf{U}^{-1}\mathbf{f}(\mathbf{\Lambda})\mathbf{U} \;\text{ и }\; \mathbf{g}(\mathbf{A}) = \mathbf{U}^{-1}\mathbf{g}(\mathbf{\Lambda})\mathbf{U}.$$

В силу того, что диагональные матрицы всегда коммутируют имеем:

$$\mathbf{f}(\mathbf{A})\mathbf{g}(\mathbf{A}) = \mathbf{U}^{-1}\mathbf{f}(\mathbf{\Lambda})\mathbf{U}\mathbf{U}^{-1}\mathbf{g}(\mathbf{\Lambda})\mathbf{U} = \mathbf{U}^{-1}\mathbf{f}(\mathbf{\Lambda})\mathbf{g}(\mathbf{\Lambda})\mathbf{U} =$$

$$= \mathbf{U}^{-1}\mathbf{g}(\mathbf{\Lambda})\mathbf{f}(\mathbf{\Lambda})\mathbf{U} = \mathbf{U}^{-1}\mathbf{g}(\mathbf{\Lambda})\mathbf{U}\mathbf{U}^{-1}\mathbf{f}(\mathbf{\Lambda})\mathbf{U} = \mathbf{g}(\mathbf{A})\mathbf{f}(\mathbf{A}).$$

Теорема 6. (Формула Кели-Гамильтона). Всякая матрица удовлетворяет своему характеристическому уравнению. Пусть $Q(\lambda) = (-1)^m \lambda^m + b_1\lambda^{m-1} + \ldots + b_m$ – характеристический полином, а $Q(\lambda) = 0$ – характеристическое уравнение. Теорема утверждает, что $Q(\mathbf{A}) = (-1)^m \mathbf{A}^m + b_1\mathbf{A}^{m-1} + \ldots + b_m\mathbf{E} \equiv \mathbf{0}$.

Доказательство. Матрица простой структуры подобна диагональной матрице: $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$. По теореме 2 о подобных матрицах $\mathbf{A}^k = \mathbf{U}\mathbf{\Lambda}^k\mathbf{U}^{-1}$. Подставим в

$$Q(\mathbf{A}) = (-1)^m \mathbf{A}^m + b_1\mathbf{A}^{m-1} + \ldots + b_m\mathbf{E}$$

вместо **A** ее выражение через $\mathbf{\Lambda}$:

$$Q(\mathbf{A}) = (-1)^m \mathbf{U\Lambda}^m\mathbf{U}^{-1} + b_1\mathbf{U\Lambda}^{m-1}\mathbf{U}^{-1} + \ldots + b_m\mathbf{E} =$$

$$= \mathbf{U}\Big((-1)^m \mathbf{\Lambda}^m + b_1\mathbf{\Lambda}^{m-1} + \ldots + b_m\mathbf{E}\Big)\mathbf{U}^{-1}.$$

В скобках стоит диагональная матрица с характеристическими полиномами на главной диагонали, в которые подставлены собственные значения, и значит тождественно равными нулю. Тогда матрица в скобках – нулевая, и теорема доказана.

<u>*Теорема 7*</u>. (Формула Лагранжа-Сильвестра). Любая функция матрицы **A**, имеющей различные собственные значения, может быть представлена в виде:

$$\mathbf{f}(\mathbf{A}) = \sum_{k=1}^{m} f(\lambda_k)\frac{(\mathbf{A}-\lambda_1\mathbf{E})\ldots(\mathbf{A}-\lambda_{k-1}\mathbf{E})(\mathbf{A}-\lambda_{k+1}\mathbf{E})\ldots(\mathbf{A}-\lambda_m\mathbf{E})}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_m)} = \sum_{k=1}^{m} f(\lambda_k)\mathbf{T}_k \qquad (6)$$

<u>*Доказательство*</u>. Представим функцию $f(x)$ в виде интерполяционного полинома Лагранжа $L_{m\text{-}1}(x)$, взяв в качестве узлов собственные значения матрицы ***A***: $\lambda_1, \lambda_2, \ldots, \lambda_m$:

$$f(x) = L_{m-1}(x) + R_{m-1}(x).$$

Подставим в эту формулу **A** вместо x:

$$\mathbf{f}(\mathbf{A}) = \mathbf{L}_{m-1}(\mathbf{A}) + \mathbf{R}_{m-1}(\mathbf{A}).$$

Остаточный член $\mathbf{R}_{m-1}(\mathbf{A})$ принимает вид

$$\mathbf{R}_{m-1}(\mathbf{A}) = \frac{f^{(m)}(\xi)}{m!}\omega(\mathbf{A}).$$

где $\omega(\mathbf{A}) = (\mathbf{A}-\lambda_1\mathbf{E})(\mathbf{A}-\lambda_2\mathbf{E})\ldots(\mathbf{A}-\lambda_{m-1}\mathbf{E})(\mathbf{A}-\lambda_m\mathbf{E})$. По теореме Кели-Гамильтона $\omega(\mathbf{A}) = \mathbf{0}$ и, следовательно, $\mathbf{f}(\mathbf{A}) = \mathbf{L}_{m-1}(\mathbf{A})$.

$$\mathbf{f}(\mathbf{A}) = \sum_{k=1}^{m} f(\lambda_k)\frac{(\mathbf{A}-\lambda_1\mathbf{E})\ldots(\mathbf{A}-\lambda_{k-1}\mathbf{E})(\mathbf{A}-\lambda_{k+1}\mathbf{E})\ldots(\mathbf{A}-\lambda_m\mathbf{E})}{(\lambda_k-\lambda_1)\ldots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\ldots(\lambda_k-\lambda_m)} =$$

$$= \sum_{k=1}^{m} f(\lambda_k)\mathrm{S}_k(\mathbf{A},\lambda_k).$$

Теорема доказана. Отметим лишь, что матричные множители $\mathrm{S}_k(\mathbf{A},\lambda_k)$ для любых матричных функций остаются одинаковыми.

Формулу Лагранжа-Сильвестра можно модифицировать и на случай кратных собственных значений, осуществив предельный переход, при котором близкие собственные значения стремятся к общему значению.

. Расширим класс рассматриваемых матричных функций.

Определение. Пусть матрица **A** не имеет кратных собственных значений, функция $f(\lambda)$ определена на спектре матрицы **A** (т.е., $f(\lambda)$ определена в точках $\lambda_1, \lambda_2, \ldots, \lambda_m$), а $L_{m\text{-}1}(x)$ – интерполяционный полином Лагранжа с узлами интерполирования λ_k. Тогда матричная функция $\mathbf{f}(\mathbf{A})$ равна $\mathbf{L}_{m-1}(\mathbf{A})$, т.е. имеет место формула (6).

В качестве упражнения проиллюстрируем некоторые из рассмотренных теорем на примере матрицы $\mathbf{A} = \begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix}$.

Характеристический полином det($\boldsymbol{A}-\lambda \boldsymbol{E}$)= $\lambda^2+101\lambda+100=0$, $\lambda_1=-100$, $\lambda_2=-1$.

Теорема 4.	λ_1	λ_2
A:	–100	–1
$e^{\mathbf{A}}$:	e^{-100}	e^{-1}
$e^{\mathbf{A}t}$:	e^{-100t}	e^{-t}
$\mathbf{A}^{-1}$:	$(-100)^{-1}$	$(-1)^{-1}$
$\mathbf{A}^n$:	$(-100)^n$	$(-1)^n$
$\mathbf{A}^T\mathbf{A}$:	(–100)(–100)	(–1)(–1)
$(\mathbf{E}\text{-}\mathbf{A})^{-1}$:	$(1-(-100))^{-1}$	$(1-(-1))^{-1}$

Теорема 6. $\mathbf{A}^2+101\mathbf{A}+100\mathbf{E}=\mathbf{0}$. Воспользуемся этой теоремой для нахождения $\mathbf{A}^{-1}$. Умножим уравнение на $\mathbf{A}^{-1}$: $\mathbf{A}+101\mathbf{E}+100\mathbf{A}^{-1}=\mathbf{0}$ или

$$\mathbf{A}^{-1}= -(\mathbf{A}+101\mathbf{E})/100$$

$$\mathbf{A}^{-1} = -\frac{1}{100}\left(\begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix} + \begin{pmatrix} 101 & 0 \\ 0 & 101 \end{pmatrix}\right) = -\frac{1}{100}\begin{pmatrix} 50 & -49 \\ -50 & 51 \end{pmatrix}$$

Теорема 7.

$$\mathbf{f}(\mathbf{A}) = f(\lambda_1)\frac{\mathbf{A}-\lambda_2\mathbf{E}}{\lambda_1-\lambda_2} + f(\lambda_2)\frac{\mathbf{A}-\lambda_1\mathbf{E}}{\lambda_2-\lambda_1},$$

$$\frac{\mathbf{A}-\lambda_2\mathbf{E}}{\lambda_1-\lambda_2} = -\frac{1}{99}\left(\begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right) = \frac{1}{99}\begin{pmatrix} 50 & 49 \\ 50 & 49 \end{pmatrix},$$

$$\frac{\mathbf{A}-\lambda_1\mathbf{E}}{\lambda_2-\lambda_1} = \frac{1}{99}\left(\begin{pmatrix} -51 & -49 \\ -50 & -50 \end{pmatrix} - \begin{pmatrix} -100 & 0 \\ 0 & -100 \end{pmatrix}\right) = \frac{1}{99}\begin{pmatrix} 49 & -49 \\ -50 & 50 \end{pmatrix},$$

$$e^{\mathbf{A}t} = e^{-100t}\frac{1}{99}\begin{pmatrix} 50 & 49 \\ 50 & 49 \end{pmatrix} + e^{-t}\frac{1}{99}\begin{pmatrix} 49 & -49 \\ -50 & 50 \end{pmatrix}.$$

Следствие 2.

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^{m} |a_{ij}| = \max(101,\ 99)=101;$$

$$\|\mathbf{A}\|_2 = \max_i \sum_{j=1}^{m} |a_{ij}| = \max(100,\ 100)=100;$$

$$\|\mathbf{A}\|_3 = \left(\sum_{i=1}^{m} \sum_{j=1}^{m} |a_{ij}|^2 \right)^{1/2} = =(51^2+2\cdot 50+49^2)^{1/2}=(10002)^{1/2}>100;$$

.

Как видно для всех норм, неравенство $\|\mathbf{A}\| \geq |\lambda|$ выполняется.

Некоторые свойства матричной экспоненты

Матричная экспонента является одной из наиболее употребительных матричных функций, поэтому обратимся к некоторым ее свойствам. Их справедливость может быть установлена проведением ряда операций с матричными рядами. Опуская строгие доказательства, ограничимся лишь указанием путей их построения.

1. $e^{\mathbf{A}} \cdot e^{\mathbf{B}} \neq e^{\mathbf{A}+\mathbf{B}}$.

Доказательство. Выпишем матричные разложения $e^{\mathbf{A}}$ и $e^{\mathbf{B}}$ и перемножим их, ограничившись членами вплоть до квадратичных.

$$e^{\mathbf{A}} \cdot e^{\mathbf{B}} = \left(\mathbf{E}+\mathbf{A}+\frac{\mathbf{A}^2}{2!}+\ldots \right) \cdot \left(\mathbf{E}+\mathbf{B}+\frac{\mathbf{B}^2}{2!}+\ldots \right) = \mathbf{E}+\mathbf{A}+\frac{\mathbf{A}^2}{2!}+\mathbf{B}+\mathbf{AB}+\frac{\mathbf{B}^2}{2!}+\ldots=$$

Теперь запишем матричный ряд для $e^{\mathbf{A}+\mathbf{B}}$:

$$e^{\mathbf{A}+\mathbf{B}} = \mathbf{E}+\mathbf{A}+\mathbf{B}+\frac{(\mathbf{A}+\mathbf{B})^2}{2!}+\ldots$$

и сравним ряды. В произведении $e^{\mathbf{A}} \cdot e^{\mathbf{B}}$ квадратичные слагаемые равны $\frac{\mathbf{A}^2}{2!}+\mathbf{AB}+\frac{\mathbf{B}^2}{2!}$, а в разложении $e^{\mathbf{A}+\mathbf{B}}$ они имеют вид $\frac{(\mathbf{A}+\mathbf{B})^2}{2!} = \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{AB}+\mathbf{BA}}{2!} + \frac{\mathbf{B}^2}{2!}$. Совпадение будет, только если **AB** = **BA**. В общем случае, когда матрицы не коммутируют $(\mathbf{AB} \neq \mathbf{BA})$, $e^{\mathbf{A}} \cdot e^{\mathbf{B}} \neq e^{\mathbf{A}+\mathbf{B}}$.

2. $e^{\mathbf{A}t} \cdot e^{\mathbf{A}q} = e^{\mathbf{A}(t+q)}$. Здесь t, q – скалярные величины.

Доказательство. Воспользуемся предыдущими результатами.

$$e^{\mathbf{A}t} \cdot e^{\mathbf{A}q} = \mathbf{E} + \mathbf{A}t + \mathbf{A}q + \frac{\mathbf{A}^2 t^2}{2!} + \mathbf{A}t\mathbf{A}q + \frac{\mathbf{A}^2 q^2}{2!} + \ldots = \mathbf{E} + \mathbf{A}(t+q) + \frac{\mathbf{A}^2 (t+q)^2}{2!} + \ldots$$

$$e^{\mathbf{A}(t+q)} = \mathbf{E} + \mathbf{A}t + \mathbf{A}q + \frac{\mathbf{A}^2 (t+q)^2}{2!} + \ldots = \mathbf{E} + \mathbf{A}(t+q) + \frac{\mathbf{A}^2 (t+q)^2}{2!} + \ldots$$

Здесь проблемы с коммутативностью матриц нет, и результат очевиден.

3. $\left(e^{\mathbf{A}}\right)^{-1} = e^{-\mathbf{A}}$.

Доказательство этого свойства состоит в проверке равенства $e^{\mathbf{A}} \cdot e^{-\mathbf{A}} = \mathbf{E}$, т.е. в непосредственном перемножении рядов для обеих матриц. Рекомендуется выполнить эти преобразования самостоятельно.

4. $\frac{d}{dt} e^{\mathbf{A}t} = \mathbf{A} e^{\mathbf{A}t} = e^{\mathbf{A}t} \mathbf{A}$.

Доказательство:

$$\frac{d}{dt} e^{\mathbf{A}t} = \frac{d}{dt}\left(\mathbf{E} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \ldots \right) = \mathbf{A} + \mathbf{A}^2 t + \frac{\mathbf{A}^3 t^2}{2!} + \ldots =$$

$$= \mathbf{A}\left(\mathbf{E} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \ldots \right) = \mathbf{A} e^{\mathbf{A}t} = \left(\mathbf{E} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \ldots \right)\mathbf{A} = e^{\mathbf{A}t}\mathbf{A}.$$

5. $\int_0^T e^{\mathbf{A}t} dt = \mathbf{A}^{-1}\left(e^{\mathbf{A}T} - \mathbf{E}\right) = \left(e^{\mathbf{A}T} - \mathbf{E}\right)\mathbf{A}^{-1}$.

Доказательство.

$$\int_0^T e^{\mathbf{A}t} dt = \int_0^T \left(\mathbf{E} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \ldots \right) dt = \left(\mathbf{E}T + \frac{\mathbf{A}T^2}{2!} + \frac{\mathbf{A}^2 T^3}{3!} + \ldots \right).$$

Умножим обе части равенства на **A** слева, а затем в правой части прибавим и вычтем **E**.

$$\mathbf{A}\int_0^T e^{\mathbf{A}t} dt = \mathbf{A}T + \frac{\mathbf{A}^2 T^2}{2!} + \frac{\mathbf{A}^3 T^3}{3!} + \ldots + \mathbf{E} - \mathbf{E} \Rightarrow \int_0^T e^{\mathbf{A}t} dt = \mathbf{A}^{-1}\left(e^{\mathbf{A}T} - \mathbf{E}\right).$$

Умножение на **A** можно выполнить и справа, и тогда:

$$\int_0^T e^{\mathbf{A}t} dt = \left(e^{\mathbf{A}T} - \mathbf{E}\right)\mathbf{A}^{-1}.$$

Выражение справа требует существования обратной матрицы **A**. Однако интеграл существует для любой матрицы. Если det(**A**)=0, можно использовать другой способ записи результата, ограничиваясь непосредственным интегрированием ряда без введения $\mathbf{A}^{-1}$.

Аналитическое решение систем линейных дифференциальных уравнений с постоянной матрицей

Рассмотрим систему обыкновенных дифференциальных уравнений первого порядка, разрешенную относительно производных

$$\frac{dx^{(i)}(t)}{dt} = f^{(i)}(t, x^{(1)}(t), x^{(2)}(t), \ldots, x^{(m)}(t)), \qquad i=1,2,\ldots,m,$$

где t – независимая переменная, $x^{(1)}(t),\ x^{(2)}(t),\ ...,\ x^{(m)}(t)$ – искомые функции, $f^{(i)}$ – функции, определенные на некотором множестве (m+1)-мерного евклидова пространства переменных $t,\ x^{(1)}(t),\ x^{(2)}(t),\ ...,\ x^{(m)}(t)$. Номер компоненты вектора здесь везде будем писать, как верхний индекс в скобках. Перейдя к векторно-матричным обозначениям

$$\mathbf{x}(t)=\begin{pmatrix} x^{(1)}(t) \\ x^{(2)}(t) \\ ... \\ x^{(m)}(t) \end{pmatrix}, \qquad \mathbf{f}(t,\mathbf{x})=\begin{pmatrix} f^{(1)}(t,x) \\ f^{(2)}(t,x) \\ ... \\ f^{(m)}(t,x) \end{pmatrix},$$

исходную систему перепишем в виде

$$\frac{d\mathbf{x}(t)}{dt}=\mathbf{f}(t,\mathbf{x}). \tag{7}$$

При этом требуется найти решение $\mathbf{x}(t)$, удовлетворяющее начальным условиям $\mathbf{x}(t_0)=\mathbf{x}_0$. Такая задача называется начальной задачей или задачей Коши.

Важным классом дифференциальных систем являются линейные системы с постоянной матрицей или постоянными коэффициентами

$$\frac{d\mathbf{x}(t)}{dt}=\mathbf{A}\mathbf{x}(t)+\mathbf{g}(t), \qquad \mathbf{x}(t_0)=\mathbf{x}_0. \tag{8}$$

Сначала обратимся к однородной системе

$$\frac{d\mathbf{x}(t)}{dt}=\mathbf{A}\mathbf{x}(t).$$

Ее решением является функция $\mathbf{x}(t)=e^{\mathbf{A}t}\mathbf{c}$, где $\mathbf{c}$ - вектор произвольных постоянных. Убедиться в этом можно непосредственной подстановкой решения в уравнение. Выражение слева $\frac{d}{dt}\left(e^{\mathbf{A}t}\mathbf{c}\right)=\mathbf{A}e^{\mathbf{A}t}\mathbf{c}$ равно выражению справа $\mathbf{A}e^{\mathbf{A}t}\mathbf{c}$.

Неоднородная система решается методом Лагранжа вариации произвольных постоянных. При этом полагаем, что элементы вектора $\mathbf{c}$ являются функциями независимой переменной $\mathbf{c}=\mathbf{c}(t)$. Подставляем искомый вид решения в уравнение

$$\mathbf{A}e^{\mathbf{A}t}\mathbf{c}(t)+e^{\mathbf{A}t}\frac{d\mathbf{c}(t)}{dt}=\mathbf{A}e^{\mathbf{A}t}\mathbf{c}(t)+\mathbf{g}(t).$$

Отсюда $e^{\mathbf{A}t}\frac{d\mathbf{c}(t)}{dt}=\mathbf{g}(t)$, и после умножения обеих частей равенства на $e^{-\mathbf{A}t}$ получаем:

$$\frac{d\mathbf{c}(t)}{dt} = e^{-\mathbf{A}t}\mathbf{g}(t).$$

Интегрируем это уравнение от t_0 до t

$$\mathbf{c}(t) - \mathbf{c}(t_0) = \int_{t_0}^{t} e^{-\mathbf{A}\tau}\mathbf{g}(\tau)d\tau,$$

и, подставив $\mathbf{c}(t)$ в искомый вид решения, определяем общее решение линейной неоднородной дифференциальной системы

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{c}(t_0) + e^{\mathbf{A}t}\int_{t_0}^{t} e^{-\mathbf{A}\tau}\mathbf{g}(\tau)d\tau.$$

Учитывая начальные условия, находим вектор $\mathbf{c}(t_0)$: $\mathbf{x}(t_0) = \mathbf{x}_0 = e^{\mathbf{A}t_0}\mathbf{c}(t_0)$ или $\mathbf{c}(t_0) = e^{-\mathbf{A}t_0}\mathbf{x}_0$ и окончательно получаем

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}_0 + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{g}(\tau)d\tau.$$

Без нарушения общности можно считать, что начальным значением независимой переменной является $t_0 = 0$. Тогда

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_{0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{g}(\tau)d\tau = e^{\mathbf{A}t}\mathbf{x}_0 + \int_{0}^{t} e^{\mathbf{A}\tau}\mathbf{g}(t-\tau)d\tau. \tag{9}$$

Последнее преобразование использует теорему о свертке:

$$\int_{0}^{t} f(t-\tau)g(\tau)d\tau = \int_{0}^{t} f(\tau)g(t-\tau)d\tau,$$

в справедливости которой легко убедиться простой заменой переменной: $t - \tau = \tau$.

Считая вектор $\mathbf{g}(t)$ постоянным $\left(\mathbf{g}(t) = \mathbf{g} = const\right)$, упростим равенство (9)

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_{0}^{t} e^{\mathbf{A}(t-\tau)}d\tau \cdot \mathbf{g} = e^{\mathbf{A}t}\mathbf{x}_0 + \left(e^{\mathbf{A}t} - \mathbf{E}\right)\mathbf{A}^{-1}\mathbf{g}.$$

Аналитическое решение систем линейных разностных уравнений с постоянной матрицей

Рассмотрим линейную систему разностных уравнений с постоянной матрицей, где k – независимая целочисленная переменная

$$\mathbf{y}(k+1) = \mathbf{B}\mathbf{y}(k) + \mathbf{g}(k). \tag{10}$$

Будем ее решать так называемым пошаговым методом, последовательно назначая в (10) значения k равными 0, 1, 2, … и обозначая $\mathbf{y}_k = \mathbf{y}(k)$

$\mathbf{y}_1 = \mathbf{B}\mathbf{y}_0 + \mathbf{g}_0$, $\mathbf{y}_2 = \mathbf{B}\mathbf{y}_1 + \mathbf{g}_1 = \mathbf{B}^2\mathbf{y}_0 + \mathbf{B}\mathbf{g}_0 + \mathbf{g}_1$,

$\mathbf{y}_3 = \mathbf{B}\mathbf{y}_2 + \mathbf{g}_2 = \mathbf{B}^3\mathbf{y}_0 + \mathbf{B}^2\mathbf{g}_0 + \mathbf{B}\mathbf{g}_1 + \mathbf{g}_2$,

…

$$\mathbf{y}_k = \mathbf{B}^k\mathbf{y}_0 + \sum_{i=0}^{k-1}\mathbf{B}^{k-i-1}\mathbf{g}_i \qquad (11)$$

В частном случае, когда $\mathbf{g}(k)=const=\mathbf{g}$, этот вектор можно вынести за знак суммы

$$\mathbf{y}_k = \mathbf{B}^k\mathbf{y}_0 + \left(\sum_{i=0}^{k-1}\mathbf{B}^i\right)\mathbf{g} = \mathbf{B}^k\mathbf{y}_0 + \left(\mathbf{B}^k - \mathbf{E}\right)\left(\mathbf{B} - \mathbf{E}\right)^{-1}\mathbf{g}. \qquad (12)$$

Последнее выражение было записано с учетом очевидного равенства.

$$\left(\mathbf{B} - \mathbf{E}\right)\left(\sum_{i=0}^{k-1}\mathbf{B}^i\right) = \left(\mathbf{B}^k - \mathbf{E}\right).$$

Устойчивость решений дифференциальных и разностных уравнений

Обратимся к системе нелинейных дифференциальных уравнений

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(t,\mathbf{x}), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad t \in [a,b], \qquad (13)$$

где t – независимая переменная, $\mathbf{x}$ – вектор решения; $\mathbf{f}(t, \mathbf{x})$ – вектор-функция, непрерывная по t и имеющая непрерывные частные производные первого порядка по компонентам вектора $\mathbf{x}$.

Большой интерес представляет исследование зависимости решения задачи Коши от начальных условий. Если незначительные изменения в $\mathbf{x}_0$ могут существенно изменить решение, то в прикладном отношении такое решение часто неприемлемо. На конечном промежутке $[a,b]$ для систем (13) с непрерывной функцией $\mathbf{f}(t, \mathbf{x})$ и свойством единственности решения имеет место *интегральная непрерывность решений*. Иными словами, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для любых двух решений $\mathbf{x}(t)$ и $\mathbf{z}(t)$ системы (13), отличающихся начальными условиями не более, чем на δ $\left(\|\mathbf{x}(t_0) - \mathbf{z}(t_0)\| < \delta\right)$, будет иметь место $\left(\|\mathbf{x}(t) - \mathbf{z}(t)\| < \varepsilon\right)$. Иначе обстоит дело для бесконечного промежутка при $t \to \infty$. Изучением этих вопросов занимается теория устойчивости.

<u>*Определение* 1</u>. Решение $\mathbf{x}(t)$ системы (13) называется *устойчивым по Ляпунову*, если для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех решений $\mathbf{z}(t)$ системы (13), удовлетворяющих неравенству $\left(\|\mathbf{x}(t_0) - \mathbf{z}(t_0)\| < \delta\right)$ справедливо неравенство $\left(\|\mathbf{x}(t) - \mathbf{z}(t)\| < \varepsilon\right)$ при всех $t \in [t_0, \infty)$.

$$\left(\forall\varepsilon > 0\right)\left(\exists\delta > 0\right)\left(\forall\mathbf{z}(t)\right)\left(\forall t \in [t_0,\infty)\right)\left(\|\mathbf{x}(t_0) - \mathbf{z}(t_0)\| < \delta \Rightarrow \|\mathbf{x}(t) - \mathbf{z}(t)\| < \varepsilon\right) \quad (14)$$

Иными словами, решение $\mathbf{x}(t)$ называется устойчивым, если другие достаточно близкие к нему в момент времени t_0 решения $\mathbf{z}(t)$ целиком находятся в узкой ε–трубке, построенной вокруг $\mathbf{x}(t)$.

Определение 2: Решение $\mathbf{x}(t)$ системы (13) называется *асимптотически устойчивым* по Ляпунову, если оно устойчиво и существует такое $\Delta > 0$, что все решения $\mathbf{z}(t)$, удовлетворяющие условию $\left(\|\mathbf{x}(t_0) - \mathbf{z}(t_0)\| < \Delta\right)$, обладают свойством

$$\lim_{t\to\infty}\|\mathbf{x}(t) - \mathbf{z}(t)\| = 0 .$$

В случае асимптотической устойчивости близкие решения не только остаются близкими друг к другу, но и неограниченно сближаются при возрастании t.

Для определения неустойчивого решения достаточно построить отрицание *определения* 1.

Для систем разностных уравнений

$$\mathbf{y}(n+1) = \mathbf{g}\big(n, \mathbf{y}(n)\big), \quad \mathbf{y}(n_0) = \mathbf{y}_0, \quad n \in [n_0, \infty) \tag{15}$$

понятие устойчивости вводится аналогично предыдущему с заменой независимой переменной t на целую переменную n. Обозначим за $\mathbf{y}(n)$ и $\mathbf{w}(n)$ два решения (15), отличающиеся начальными условиями $\mathbf{y}(n_0)$ и $\mathbf{w}(n_0)$.

Определение 3: Решение $\mathbf{y}(n)$ называется *устойчивым*, если для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всех решений $\mathbf{w}(n)$ системы (15), удовлетворяющих неравенству $\left(\|\mathbf{y}(n_0) - \mathbf{w}(n_0)\| < \delta\right)$ справедливо неравенство $\left(\|\mathbf{y}(n) - \mathbf{w}(n)\| < \varepsilon\right)$ при всех $n \in [n_0, \infty)$

$$\big(\forall\varepsilon > 0\big)\big(\exists\delta > 0\big)\big(\forall\mathbf{w}(n)\big)\big(\forall n \in [n_0, \infty)\big)\left(\|\mathbf{y}(n_0) - \mathbf{w}(n_0)\| < \delta \Rightarrow \|\mathbf{y}(n) - \mathbf{w}(n)\| < \varepsilon\right) \tag{16}$$

Понятие *асимптотической устойчивости* предлагается определить самостоятельно на основе *определения 2*.

Сформулированные определения позволяют сделать суждение об устойчивости после анализа уже полученных решений. С практической точки зрения важно судить об устойчивости, не решая задачу. Это возможно, в частности, для линейных систем с постоянной матрицей (8)

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{g}(t), \qquad \mathbf{x}(t_0) = \mathbf{x}_0 .$$

Будем называть их устойчивыми (асимптотически устойчивыми, неустойчивыми), если все их решения устойчивы (асимптотически устойчивы, неустойчивы).

Пусть $\mathbf{x}(t)$ и $\mathbf{z}(t)$ – два различных решения (8), отличающиеся начальными условиями. В соответствии с (9) они имеют вид

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{g}(\tau)d\tau = e^{\mathbf{A}t}\mathbf{x}_0 + \int_0^t e^{\mathbf{A}\tau}\mathbf{g}(t-\tau)d\tau ,$$

$$\mathbf{z}(t) = e^{\mathbf{A}t}\mathbf{z}_0 + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{g}(\tau)d\tau = e^{\mathbf{A}t}\mathbf{z}_0 + \int_0^t e^{\mathbf{A}\tau}\mathbf{g}(t-\tau)d\tau .$$

Вычтем из первой формулы вторую. После сокращения интегралов получаем

$$\mathbf{x}(t) - \mathbf{z}(t) = e^{\mathbf{A}t}\left(\mathbf{x}_0 - \mathbf{z}_0\right).$$

Пусть первоначально собственные значения матрицы $\mathbf{A}$ различны. Тогда, используя для матричной экспоненты формулу Лагранжа-Сильвестра (6), имеем

$$\mathbf{x}(t) - \mathbf{z}(t) = e^{\mathbf{A}t}\left(\mathbf{x}_0 - \mathbf{z}_0\right) = \sum_{k=1}^{m} e^{\lambda_k t}\mathbf{T}_k\left(\mathbf{x}_0 - \mathbf{z}_0\right).$$

Обращаясь к *определениям* 1 и 2, приходим к выводу о том, что для обеспечения неравенства $\left(\|\mathbf{x}(t) - \mathbf{z}(t)\| < \varepsilon\right)$ элементы матричной экспоненты при $t \to \infty$ должны быть ограничены. А это, в свою очередь, требует, чтобы вещественные части $\mathrm{Re}(\lambda_k)$ собственных значений были бы неположительные. Для асимптотической устойчивости условие $\lim\limits_{t\to\infty}\|\mathbf{x}(t) - \mathbf{z}(t)\| = 0$ выполняется, когда элементы матричной экспоненты при $t \to \infty$ стремятся к нулю, а вещественные части собственных значений соответственно отрицательные.

Если среди собственных значений есть кратные, условия несколько корректируются. Пусть, например, собственное значение λ_k имеет кратность s. Тогда в решении этой группе собственных значений отвечает слагаемое $P_{s-1}(t)e^{\lambda_k t}$, где $P_{s-1}(t)$ – полином степени s-1. Если для $\mathrm{Re}(\lambda_k) < 0$ асимптотическая устойчивость обеспечивается независимо от кратности корня ($P_{s-1}(t)e^{\lambda_k t} \to 0$ при $t \to \infty$), то при нулевой вещественной части $P_{s-1}(t) \to \pm\infty$ при $t \to \infty$ и условие $\left(\|\mathbf{x}(t) - \mathbf{z}(t)\| < \varepsilon\right)$ не выполняется.

Подведем итоги.

1. Для асимптотической устойчивости необходимо и достаточно, чтобы для всех собственных значений выполнялись условия $\mathrm{Re}(\lambda_k) < 0$.

2. Для устойчивости необходимо, чтобы $\mathrm{Re}(\lambda_k) \le 0$. При этом достаточно, чтобы среди собственных значений с нулевой вещественной частью не было бы кратных.

3. Для неустойчивости необходимо наличие хотя бы одного собственного значения с $\mathrm{Re}(\lambda_k) > 0$ или кратных собственных значений с $\mathrm{Re}(\lambda_k) = 0$.

Теперь обратимся к системе разностных уравнений с постоянной матрицей

$$\mathbf{y}(n+1) = \mathbf{B}\mathbf{y}(n) + \mathbf{g}(n).$$

Пусть $\mathbf{y}(n)$ и $\mathbf{w}(n)$ – ее два различных решения, отличающиеся начальными условиями. В соответствии с (11) они имеют вид

$$\mathbf{y}(n) = \mathbf{B}^n\mathbf{y}(0) + \sum_{k=0}^{n-1}\mathbf{B}^{n-k-1}\mathbf{g}(k) = \mathbf{B}^n\mathbf{y}(0) + \sum_{k=0}^{n-1}\mathbf{B}^k\mathbf{g}(n-k-1),$$

$$\mathbf{w}(n) = \mathbf{B}^n\mathbf{w}(0) + \sum_{k=0}^{n-1}\mathbf{B}^{n-k-1}\mathbf{g}(k) = \mathbf{B}^n\mathbf{w}(0) + \sum_{k=0}^{n-1}\mathbf{B}^k\mathbf{g}(n-k-1).$$

Вычитая из первой формулы вторую, после сокращения сумм получаем

$$\mathbf{y}(n) - \mathbf{w}(n) = \mathbf{B}^n\left(\mathbf{y}_0 - \mathbf{w}_0\right).$$

Если все собственные значения μ_k матрицы $\mathbf{B}$ различны, то, воспользовавшись формулой Лагранжа-Сильвестра для $\mathbf{B}^n$, имеем

$$\mathbf{y}(n) - \mathbf{w}(n) = \sum_{k=1}^{m}\mu_k^n\mathbf{T}_k\left(\mathbf{y}_0 - \mathbf{w}_0\right),$$

где μ_k – собственные значения матрицы $\mathbf{B}$. Аналогично предыдущему для обеспечения неравенства $\left(\|\mathbf{y}(n) - \mathbf{w}(n)\| < \varepsilon\right)$ элементы матрицы $\mathbf{B}^n$ при $n \to \infty$ должны быть ограничены. А это, в свою очередь, требует выполнения условий $|\mu_k| \le 1$ для всех собственных значений. Для асимптотической устойчивости неравенства должны быть строгими: $|\mu_k| < 1$. Если собственное значение μ_k имеет кратность s, то, как и для дифференциальных уравнений, в решении появляется слагаемое $P_{s-1}(n)\cdot\mu_k^n$, где $P_{s-1}(n)$ – полином степени s-1. Для $|\mu_k| < 1$ этот факт не оказывает влияния на условие устойчивости, но для $|\mu_k| = 1$ условие устойчивости нарушается, если $P_{s-1}(n) \to \pm\infty$ при $n \to \infty$. Как результат, сформулируем условия устойчивости.

1. Для асимптотической устойчивости необходимо и достаточно, чтобы для всех собственных значений выполнялись условия $|\mu_k| < 1$.

2. Для устойчивости необходимо, чтобы $|\mu_k| \le 1$. При этом достаточно, чтобы среди собственных значений с единичными модулями не было бы кратных.

3. Для неустойчивости необходимо наличие хотя бы одного собственного значения с $|\mu_k| > 1$ или кратных собственных значений с $|\mu_k| = 1$.

Глубже осознать понятие устойчивости позволяет ответ на следующий вопрос.

Вопрос. *В какой взаимосвязи находятся понятия «устойчивости» и «ограниченности»? Имеется четыре варианта ответа.*

1. Из устойчивости решения следует его ограниченность.

2. Из ограниченности решения следует его устойчивость.

3. Оба предыдущих факта имеют место и по сути свойства ограниченности и устойчивости весьма близки.

4. Это различные понятия. Из ограниченности решения еще не следует его устойчивость и из устойчивости не следует ограниченность.

Рекомендуется хорошо подумать над ответом прежде, чем переходить к следующему примеру, частично отвечающему на этот вопрос.

Пример. $$\frac{dx}{dt} = -x + t^2 + 2t$$

Рассмотрим первый вариант: «*Следует ли из устойчивости решения его ограниченность?*».

Уравнение первого порядка имеет вид (8) с матрицей 1×1 и собственным значением, равным -1. Асимптотическая устойчивость очевидна. Между тем, решение *x(t)* и другое решение *y(t)*, отличающееся начальным условием, имеют вид:

$$x(t) = e^{-t}x_0 + t^2,$$

$$y(t) = e^{-t}y_0 + t^2.$$

Разность двух решений, как и положено, стремится к нулю при $t \to \infty$, в то время как оба решения не ограничены и стремятся к t^2 . Никакая ограниченность не наблюдается!