Введем в рассмотрение фундаментальную матрицу $U=\|u_{ik}\|_1^n$ для матрицы A (в k-м столбце матрицы U стоят координаты k-го собственного вектора $\overset{k}{u}$ матрицы A; $k=1, 2, \ldots, n$). Тогда (см. гл. III, стр. 86) характеристическому числу $\lambda_1\lambda_2 \ldots \lambda_p$ матрицы $\mathfrak{A}_p$ будет соответствовать собственный вектор с координатами

$$U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} \quad (1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n). \tag{129}$$

По теореме Фробениуса все числа (129) отличны от нуля и одного знака. Умножая векторы $\overset{1}{u}, \overset{2}{u}, \ldots, \overset{n}{u}$ на ± 1, можно сделать все миноры (129) положительными:

$$U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} > 0 \quad \begin{pmatrix} 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n; \\ p=1, 2, \ldots, n \end{pmatrix}. \tag{130}$$

Фундаментальная матрица $U=\|u_{ik}\|_1^n$ для матрицы A связана с матрицей A равенством

$$A = U\{\lambda_1, \lambda_2, \ldots, \lambda_n\}U^{-1}. \tag{131}$$

Но тогда

$$A' = U'^{-1}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}U'. \tag{132}$$

Сопоставляя (131) с (132), мы видим, что матрица

$$V = U'^{-1} \tag{133}$$

является фундаментальной для транспонированной матрицы A' при тех же характеристических числах $\lambda_1, \lambda_2, \ldots, \lambda_n$. Но из осцилляционности матрицы A следует осцилляционность транспонированной матрицы A'. Поэтому и для матрицы V при любом $p=1, 2, \ldots, n$ все миноры

$$V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} \quad (1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n) \tag{134}$$

отличны от нуля и имеют один и тот же знак.

С другой стороны, согласно (133) матрицы U и V связаны равенством

$$U'V = E.$$

Переходя к p-м ассоциированным матрицам (см. гл. I, § 4), будем иметь:

$$\mathfrak{U}_p\mathfrak{V}_p = \mathfrak{E}_p.$$

Отсюда, в частности, записывая, что диагональный элемент матрицы $\mathfrak{E}_p$ равен единице, получим:

$$\sum_{1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n} U\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} = 1. \tag{135}$$

В левой части этого равенства первые множители в слагаемых положительны, а вторые — отличны от нуля и одного знака. Тогда очевидно, что и вторые сомножители положительны, т. е.

$$V\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ 1 & 2 & \ldots & p \end{pmatrix} > 0 \quad \begin{pmatrix} 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n; \\ p=1, 2, \ldots, n \end{pmatrix}. \tag{136}$$

Таким образом, для матриц $U=\|u_{ik}\|_1^n$ и $V=U'^{-1}$ одновременно имеют место неравенства (130) и (136).

Выражая миноры матрицы V через миноры обратной матрицы $V^{-1}=U'$ по известным формулам [см. стр. 31], получим:

$$V\begin{pmatrix} j_1\, j_2 \cdots j_{n-p} \\ 1\ \ 2 \ldots n-p \end{pmatrix} = \frac{(-1)^{np+\sum\limits_{\nu=1}^{p} i_\nu}}{|U|} U \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ n & n-1 & \ldots & n-p+1 \end{pmatrix}, \tag{137}$$

где $i_1<i_2<\ldots<i_p$ и $j_1<j_2<\ldots<j_{n-p}$ вместе дают полную систему индексов $1, 2, \ldots, n$. Так как в силу (130) $|U|>0$, то из (136) и (137) вытекает:

$$(-1)^{np+\sum\limits_{\nu=1}^{p} i_\nu} U \begin{pmatrix} i_1\, i_2 \ldots i_p \\ 1\ \ 2 \ldots p \end{pmatrix} > 0 \quad \begin{pmatrix} 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n; \\ p=1, 2, \ldots, n \end{pmatrix}. \tag{138}$$

Пусть теперь $u=\sum\limits_{k=g}^{h} c_k \overset{k}{u} \left(\sum\limits_{k=g}^{h} c_k^2>0\right)$. Мы покажем, что из неравенств (130) следует вторая часть неравенства (128):

$$S_u^+ \leqslant h-1, \tag{139}$$

а из неравенств (138) — первая:

$$S_u^- \geqslant g-1. \tag{140}$$

Допустим, что $S_u^+>h-1$. Тогда можно указать такие $h+1$ координат вектора u

$$u_{i_1},\ u_{i_2},\ \ldots,\ u_{i_{h+1}} \ (1 \leqslant i_1 < i_2 < \ldots < i_{h+1} \leqslant n), \tag{141}$$

что

$$u_{i_\alpha} u_{i_{\alpha+1}} \leqslant 0 \quad (\alpha=1,\ 2,\ \ldots,\ h).$$

При этом координаты (141) не могут все одновременно равняться нулю, так как тогда, приравнивая нулю соответствующие координаты вектора $u=\sum\limits_{k=1}^{h} c_k \overset{k}{u}$ $\left(c_1=\ldots=c_{g-1}=0;\ \sum\limits_{k=1}^{h} c_k^2>0\right)$, мы получили бы систему однородных уравнений

$$\sum_{k=1}^{h} c_k u_{i_\alpha}{}^k = 0 \quad (\alpha=1,\ 2,\ \ldots,\ h)$$

с ненулевым решением $c_1, c_2, \ldots, c_h$; в то же время определитель этой системы

$$U\begin{pmatrix} i_1\, i_2 \ldots i_h \\ 1\ \ 2 \ldots h \end{pmatrix}$$

согласно (130) отличен от нуля.

Рассмотрим теперь равный нулю определитель

$$\begin{vmatrix} u_{i_1 1} & \ldots & u_{i_1 h} & u_{i_1} \\ u_{i_2 1} & \ldots & u_{i_2 h} & u_{i_2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ u_{i_{h+1} 1} & \ldots & u_{i_{h+1} h} & u_{i_{h+1}} \end{vmatrix} = 0.$$

Раскроем его по элементам последней вертикали:

$$\sum_{\alpha=1}^{h+1} (-1)^{h+\alpha+1} u_{i_\alpha} U \begin{pmatrix} i_1 \ldots i_{\alpha-1}\ i_{\alpha+1} \ldots i_{h+1}; \\ 1 \ldots\ldots\ldots\ldots h \end{pmatrix} = 0.$$

Но такое равенство не может иметь места, так как в левой части нет двух слагаемых разных знаков и по крайней мере одно слагаемое отлично от нуля. Таким образом, допущение $S_u^+ > h-1$ привело нас к противоречию, и неравенство (139) можно считать установленным.

Введем в рассмотрение векторы

$$\overset{k}{u}{}^* = (u_{1k}^*,\ u_{2k}^*,\ \ldots,\ u_{nk}^*) \qquad (k=1,\ 2,\ \ldots,\ n),$$

где

$$u_{ik}^* = (-1)^{n+i+k} u_{ik} \qquad (i,\ k=1,\ 2,\ \ldots,\ n);$$

тогда для матрицы $U^* = \| u_{ik}^* \|_1^n$ в силу (138) будем иметь:

$$U^* \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ n & n-1 & \ldots & n-p+1 \end{pmatrix} > 0 \quad \begin{pmatrix} 1 \leqslant i_1 < i_2 < \ldots < i_p \leqslant n; \\ p=1,\ 2,\ \ldots,\ n \end{pmatrix}. \tag{142}$$

Но неравенства (142) аналогичны неравенствам (130). Поэтому, полагая

$$u^* = \sum_{k=g}^{h} (-1)^k c_k \overset{k}{u}{}^*, \tag{143}$$

будем иметь неравенство, аналогичное неравенству (139)[1]):

$$S_{u^*}^+ \leqslant n-g. \tag{144}$$

Пусть $u=(u_1,\ u_2,\ \ldots,\ u_n)$, а $u^*=(u_1^*,\ u_2^*,\ \ldots,\ u_n^*)$. Легко видеть, что

$$u_i^* = (-1)^i u_i \qquad (i=1,\ 2,\ \ldots,\ n).$$

Поэтому

$$S_{u^*}^+ + S_u^- = n-1$$

и, следовательно, в силу (144) имеет место соотношение (140).

Неравенство (128) установлено. Поскольку из него получается при $g=h=k$ утверждение 2 теоремы, то теорема доказана полностью.

3. Рассмотрим применение доказанной теоремы к исследованию малых колебаний n масс $m_1,\ m_2,\ \ldots,\ m_n$, сосредоточенных в n подвижных точках $x_1 < x_2 < \ldots < x_n$ сегментного упругого континуума (струна или стержень конечной длины), простирающегося (в состоянии равновесия) вдоль отрезка $0 \leqslant x \leqslant l$ оси x.

Обозначим через $K(x,\ s)$ $(0 \leqslant x,\ s \leqslant l)$ функцию влияния этого континуума [$K(x,\ s)$ — прогиб в точке x под действием единичной силы, приложенной в точке s], а через k_{ij} — коэффициенты влияния для данных n масс:

$$k_{ij} = K(x_i,\ x_j) \qquad (i,\ j=1,\ 2,\ \ldots,\ n).$$

---

[1]) В неравенствах (142) в векторы $\overset{k}{u}$ $(k=1,\ 2,\ \ldots,\ n)$ идут в обратном порядке $\overset{n}{u},\ \overset{n-1}{u},\ \ldots$ Вектору $\overset{g}{u}$ предшествует $n-g$ векторов этого ряда.

Если в точках $x_1, x_2, \ldots, x_n$ приложены n сил $F_1, F_2, \ldots, F_n$, то соответствующий статический прогиб $y(x)$ $(0 \leqslant x \leqslant l)$ в силу линейного наложения прогибов выразится формулой

$$y(x) = \sum_{j=1}^{n} K(x, x_j) F_j.$$

Заменяя здесь силы F_j силами инерции $-m_j \frac{\partial^2}{\partial t^2} y(x_j; t)$ $(j = 1, 2, \ldots, n)$, получим уравнение свободных колебаний

$$y(x) = -\sum_{j=1}^{n} m_j K(x, x_j) \frac{\partial^2}{\partial t^2} y(x_j; t). \tag{145}$$

Будем искать гармонические колебания континуума в виде

$$y(x) = u(x) \sin(\omega t + \alpha) \qquad (0 \leqslant x \leqslant l). \tag{146}$$

Здесь $u(x)$ — амплитудная функция, ω — частота, α — начальная фаза. Подставляя это выражение для $y(x)$ в (145) и сокращая на $\sin(\omega t + \alpha)$, получим:

$$u(x) = \omega^2 \sum_{j=1}^{n} m_j K(x, x_j) u(x_j). \tag{147}$$

Введем обозначения для переменных прогибов и для амплитудных прогибов в точках расположения масс:

$$y_i = y(x_i, t), \quad u_i = u(x_i) \qquad (i = 1, 2, \ldots, n).$$

Тогда

$$y_i = u_i \sin(\omega t + \alpha) \qquad (i = 1, 2, \ldots, n).$$

Введем еще *приведенные амплитудные прогибы и приведенные коэффициенты влияния*

$$\tilde{u}_i = \sqrt{m_i}\, u_i, \quad a_{ij} = \sqrt{m_i m_j}\, k_{ij} \qquad (i, j = 1, 2, \ldots, n). \tag{148}$$

Заменяя в (147) x последовательно на x_i $(i = 1, 2, \ldots, n)$, получим систему уравнений для амплитудных прогибов:

$$\sum_{j=1}^{n} a_{ij} \tilde{u}_j = \lambda \tilde{u}_i \qquad \left(\lambda = \frac{1}{\omega^2};\ i = 1, 2, \ldots, n\right). \tag{149}$$

Отсюда видно, что амплитудный вектор $\tilde{u} = (\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_n)$ есть собственный вектор матрицы $A = \| a_{ij} \|_1^n = \| \sqrt{m_i m_j}\, k_{ij} \|_1^n$ при $\lambda = \frac{1}{\omega^2}$ (ср. с гл. X, § 8).

В результате подробного анализа[1]) устанавливается, что *матрица коэффициентов влияния* $\| k_{ij} \|_1^n$ *сегментного континуума всегда является осцилляционной матрицей.* Но тогда и матрица $A = \| a_{ij} \|_1^n = \| \sqrt{m_i m_j} k_{ij} \|_1^n$ является осцилляционной! Поэтому матрица A (согласно теореме 13) имеет n положительных характеристических чисел

$$\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0,$$

---

[1]) См. [10д, е] и [7], гл. III.

т. е. существует n гармонических колебаний континуума с *различными* частотами:

$$(0<)\,\omega_1<\omega_2<\ldots<\omega_n \qquad \left(\lambda_i=\frac{1}{\omega_i^2};\quad i=1,\ 2,\ \ldots,\ n\right).$$

В силу той же теоремы основному тону с частотой ω_1 соответствуют амплитудные прогибы, отличные от нуля и одного знака. В ряду амплитудных прогибов, отвечающих первому обертону с частотой ω_2, имеется точно одна перемена знака и вообще в ряду амплитудных прогибов для обертона с частотой ω_j имеется точно $j-1$ перемен знака ($j=1,\ 2,\ \ldots,\ n$).

Из того факта, что матрица коэффициентов влияния $\|k_{ij}\|_1^n$ осцилляционна, вытекают и другие осцилляционные свойства континуума: 1) при $\omega=\omega_1$ амплитудная функция $u(x)$, связанная с амплитудными прогибами формулой (147), не имеет узлов; и вообще при $\omega=\omega_j$ эта функция имеет $j-1$ узлов ($j=1,\ 2,\ \ldots,\ n$); 2) узлы двух смежных тонов перемежаются и т. д.

На обосновании этих свойств мы не можем здесь останавливаться[1]).

[1]) См. [7], гл. III, IV.

ГЛАВА XIV

РАЗЛИЧНЫЕ КРИТЕРИИ РЕГУЛЯРНОСТИ И ЛОКАЛИЗАЦИЯ СОБСТВЕННЫХ ЗНАЧЕНИЙ

§ 1. Критерий регулярности Адамара и его обобщения

Пусть $A = \| a_{ik} \|_1^n$ — произвольная $n \times n$-матрица с комплексными элементами. Допустим, что эта матрица вырождена, т. е. $|A| = 0$. Тогда существуют такие числа $x_1, x_2, \ldots, x_n$ с максимальным $|x_k| > 0$, что[1])

$$\sum_{j=1}^{n} a_{kj} x_j = 0. \tag{1}$$

Но тогда

$$|a_{kk}|\,|x_k| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|\,|x_j| \leqslant |x_k| \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|.$$

Сокращая на $|x_k|$, получаем:

$$|a_{kk}| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|. \tag{2}$$

Поэтому, если выполняются *условия Адамара*

$$H_i \equiv |a_{ii}| - \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| > 0 \qquad (i = 1, \ldots, n), \tag{3}$$

то неравенство типа (2) невозможно и, следовательно, матрица A является *регулярной* (невырожденной), т. е. $|A| \neq 0$.

Таким образом, справедлива следующая

Теорема 1 (Адамара). *Если для матрицы $A = \| a_{ik} \|_1^n$ выполняются n неравенств* (3), *то матрица A является невырожденной.*

Условие $H_i > 0$ означает, что модуль диагонального элемента a_{ii} превосходит (строго!) сумму модулей всех остальных элементов i-й строки. Такой элемент a_{ii} называется доминирующим (для своей строчки). Условия Адамара требуют, чтобы все диагональные элементы матрицы A были *доминирующими* (для своих строк).

Замечание 1. Если выполняются условия Адамара (3), то для $\operatorname{mod}|A|$ справедлива следующая оценка снизу:

$$\operatorname{mod}|A| \geqslant H_1 H_2 \ldots H_n > 0. \tag{4}$$

---

[1]) Равенства (1) выполняются при любом $k = 1, \ldots, n$. Мы же берем только то значение k, при котором $|x_k|$ является максимальным.

Для того чтобы убедиться в справедливости неравенства (4), введем вспомогательную матрицу $F = \| f_{ij} \|_1^n$, где

$$f_{ij} = \frac{a_{ij}}{H_i} \qquad (i,\ j = 1,\ \ldots,\ n), \tag{5}$$

для которой, очевидно,

$$|f_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^{n} |f_{ij}| = 1 \qquad (i = 1,\ \ldots,\ n). \tag{6}$$

Обозначим через λ_0 какое-либо характеристическое число этой матрицы. Числу λ_0 соответствует собственный вектор $(x_1,\ x_2,\ \ldots,\ x_n)$ с максимальным $|x_k| > 0$. Тогда

$$\lambda_0 x_k = \sum_{j=1}^{n} f_{kj} x_j. \tag{7}$$

Из этого равенства с учетом соотношений (6) получаем

$$|\lambda_0|\,|x_k| \geqslant |f_{kk}|\,|x_k| - \sum_{\substack{j=1 \\ j \neq k}}^{n} |f_{kj}|\,|x_j| \geqslant |x_k| \Big(|f_{kk}| - \sum_{\substack{j=1 \\ j \neq k}}^{n} |f_{kj}| \Big) = |x_k|.$$

Сокращая на $|x_k|$, найдем:

$$|\lambda_0| \geqslant 1.$$

Но определитель $|F|$ равен произведению характеристических чисел матрицы F. Каждое из них по модулю $\geqslant 1$. Поэтому и

$$\operatorname{mod} |F| \geqslant 1. \tag{8}$$

С другой же стороны,

$$|F| = \frac{|A|}{H_1 H_2 \ldots H_n}. \tag{9}$$

Из (8) и (9) сразу следует искомое неравенство (4).

Заметим еще, что для всего класса матриц, удовлетворяющих условиям Адамара с заданными значениями $H_1,\ \ldots,\ H_n$, оценка (4) не может быть улучшена, так как неравенство (4) переходит в равенство, если в качестве матрицы A взять матрицу $\| H_i \delta_{ik} \|$.

Замечание 2. Поскольку $|A| = |A'|$, то, заменяя матрицу A транспонированной матрицей A', получаем достаточные условия невырожденности матрицы A в виде условий Адамара для столбцов

$$G_i \equiv |a_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}| > 0 \qquad (i = 1,\ \ldots,\ n). \tag{10}$$

При выполнении этих условий вместо (4) будем иметь

$$\operatorname{mod} |A| \geqslant G_1 G_2 \ldots G_n. \tag{11}$$

Пусть C — произвольная невырожденная $n \times n$-матрица. Тогда матрицы A и AC одновременно являются невырожденными. Поэтому в условиях (3), (10), а также в оценках (4) и (11) можно матрицу A заменить на AC. Варьируя матрицу C, будем получать различные (неэквивалентные между собой) достаточные условия невырожденности, а также оценки для $|A|$, аналогичные (4) и (11). В частности, за счет подбора надлежащей матрицы C можно осуществить произвольную

перестановку столбцов. Тогда вместо условий (3) получим условия

$$H'_i \equiv |a_{i\mu_i}| - \sum_{\substack{j=1\\ j\neq\mu_i}}^{n} |a_{ij}| > 0 \qquad (i=1, \ldots, n), \tag{12}$$

где $(\mu_1, \ldots, \mu_n)$ — фиксированная, но произвольная перестановка индексов $1, 2, \ldots, n$.

Другими словами, матрица $A = \|a_{ij}\|_n^1$ будет невырожденной, если в каждой ее строке имеется доминирующий (не обязательно диагональный!) элемент и эти n доминирующих элементов расположены в различных столбцах.

Аналогичное предложение имеет место для столбцов.

Пусть теперь выполняются ослабленные условия Адамара

$$H_i = |a_{ii}| - \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \geqslant 0 \qquad (i=1, \ldots, n). \tag{13}$$

В этом случае каждый диагональный элемент является *слабо доминирующим* для своей строки.

Допустим, что матрица A вырождена и $Ax=0$, вектор-столбец $x=(x_1, \ldots, x_n)\neq 0$ имеет ровно p элементов x_k с максимальным модулем $|x_k|$, и пусть сначала $p<n$. Перенумеруем координаты вектора x так, чтобы этими максимальными по модулю были первые p координат:

$$|x_1| = \ldots = |x_p| > |x_j| \qquad (j=p+1, \ldots, n).$$

При этом равенство $Ax=0$ сохранится, если мы совершим некоторую (но одну и ту же) перестановку строк и столбцов матрицы A. После этого можно написать:

$$a_{kk}x_k = -\sum_{\substack{j=1\\ j\neq k}}^{n} a_{kj}x_j \qquad (k=1, \ldots, p),$$

откуда

$$|a_{kk}|\,|x_k| \leqslant \Big(\sum_{\substack{j=1\\ j\neq k}}^{p} |a_{kj}|\Big)|x_k| + \sum_{j=p+1}^{n} |a_{kj}|\,|x_j| \leqslant$$

$$\leqslant |x_k| \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}| \qquad (k=1, \ldots, n). \tag{14}$$

Сокращая на $|x_k|$, получим

$$|a_{kk}| \leqslant \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}| \qquad (k=1, \ldots, n). \tag{15}$$

Сопоставляя эти соотношения с ослабленными условиями Адамара (13), которые имеют место по условию, заключаем, что во всех соотношениях (15), а значит и в (14), имеет место знак равенства. А это возможно лишь тогда, когда

$$\sum_{j=p+1}^{n} |a_{kj}| = 0 \qquad (k=p+1, \ldots, n),$$

т. е. матрица A имеет вид

$$A=\begin{pmatrix} A_1 & \overset{p}{\overline{0}} \\ A_3 & A_4 \end{pmatrix} \Big\}\, p. \tag{16}$$

Но матрица, которая одной и той же перестановкой строк и столбцов приводится к виду (16), называется разложимой (см. гл. 13, § 1). Таким образом, при $p<n$ A — разложимая матрица.

Если же $p=n$, то во всех соотношениях (15) и, следовательно, во всех n ослабленных условиях Адамара (13) имеет место знак равенства.

К этим выводам мы пришли, допустив, что A — вырожденная матрица.

Таким образом, нами доказана следующая теорема, представляющая собой уточнение теоремы Адамара.

Теорема 2 (Ольги Тауски). *Если для неразложимой матрицы A выполняются ослабленные условия Адамара* (13) *и по крайней мере в одном из этих условий имеет место знак* $>$, *то матрица A невырождена.* Само собой разумеется, что и в этой теореме условия $H_i \geqslant 0$ $(i=1, \ldots, n)$ могут быть заменены условиями $G_i \geqslant 0$ $(i=1, \ldots, n)$.

Для дальнейших обобщений теоремы Адамара нам понадобится понятие о норме прямоугольной матрицы. Этому понятию посвящается следующий параграф.

§ 2. Норма матрицы

В n-мерном пространстве R векторов-столбцов x введем понятие о норме вектора. Каждому вектору $x \in R$ ставим в соответствие некоторое вещественное неотрицательное число $\|x\|_R$ или просто $\|x\|$ так, чтобы для произвольных векторов x, y из R и произвольного скаляра λ выполнялись следующие условия:

1° $$\|x+y\| \leqslant \|x\|+\|y\|,$$

2° $$\|\lambda x\| = |\lambda|\,\|x\|,$$

3° $$\|x\| > 0, \text{ если } x \neq 0.$$

Полагая в 2° $\lambda=0$, получим, что $\|x\|=0$, если $x=0$. Кроме того, из 2° сразу следует для любых векторов x, $y \in R$: $\|x-y\| \geqslant \|x\|-\|y\|$.

Так, например, можно ввести «кубическую» норму вектора

$$\|x\|_{\mathrm{I}} = \max_{1 \leqslant i \leqslant n} |x_i| \tag{17}$$

или «октаэдрическую» норму

$$\|x\|_{\mathrm{II}} = \sum_{i=1}^{n} |x_i|. \tag{17'}$$

«Эрмитову» (в случае вещественного пространства R «евклидову») норму $\|x\|_{\mathrm{III}}$ определяют равенством[1])

$$\|x\|_{\mathrm{III}} = \sqrt{\sum_{i=1}^{n} |x_i|^2}. \tag{17''}$$

[1]) См. гл. IX, § 2.

Легко проверяется, что все эти нормы удовлетворяют постулатам 1°, 2° и 3°.

Рассмотрим теперь произвольную прямоугольную $m \times n$-матрицу A и связанное с нею линейное преобразование $y = Ax$. x — n-мерный вектор-столбец из n-мерного пространства R, а y — m-мерный вектор-столбец из m-мерного пространства S.

Введем в этих пространствах нормы векторов $\|x\|_R = \|x\|$ и $\|y\|_S = \|y\|$. После этого норму прямоугольной матрицы A определим равенством

$$\|A\| = \sup_{\substack{x \in R \\ x \neq 0}} \frac{\|Ax\|_S}{\|x\|_R}. \tag{18}$$

Норма $m \times n$-матрицы A определяется как самой матрицей A, так и теми векторными нормами, которые введены в пространствах R и S. При изменении этих норм изменяется и норма матрицы.

Из определения нормы следует очевидное соотношение

$$\|Ax\|_S \leqslant \|A\| \, \|x\|_R. \tag{18'}$$

Для двух $m \times n$-матриц A и B при одном и том же определении векторных норм имеем соотношение

$$\|A + B\| \leqslant \|A\| + \|B\|. \tag{19}$$

Кроме того, очевидно, что

$$\|\lambda A\| = |\lambda| \, \|A\|. \tag{19'}$$

Пусть $p \times n$-матрица B отображает n-мерное пространство R в p-мерное S, а $m \times p$-матрица A отображает p-мерное пространство S в m-мерное T. Очевидно, матрица AB отображает R в T. Вводя в пространствах R, S и T векторные нормы и определяя с их помощью нормы матриц $\|A\|$, $\|B\|$, $\|AB\|$, легко приходим к неравенству

$$\|AB\| \leqslant \|A\| \, \|B\|. \tag{19''}$$

Так, например, если исходить из «кубических» векторных норм $\|x\|_{\mathrm{I}} = \max\limits_{1 \leqslant k \leqslant n} |x_k|$, $\|y\|_{\mathrm{I}} = \max\limits_{1 \leqslant i \leqslant m} |y_i|$, то норма матрицы $A = \|a_{ik}\|$ $(i = 1, \ldots, m;\ k = 1, \ldots, n)$ определяется формулой

$$\|A\| = \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|. \tag{20}$$

Действительно, в этом случае

$$\|Ax\|_{\mathrm{I}} = \max_{1 \leqslant i \leqslant m} \left| \sum_{k=1}^{n} a_{ik} x_k \right| \leqslant \max_{1 \leqslant k \leqslant n} |x_k| \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|,$$

и потому

$$\frac{\|Ax\|_{\mathrm{I}}}{\|x\|} \leqslant \max_{1 \leqslant i \leqslant m} \sum_{k=1}^{n} |a_{ik}|.$$

В то же время здесь знак $=$ имеет место, если выбрать координаты $x_1, x_2, \ldots, x_n$ вектора x так, чтобы $|x_1| = |x_2| = \ldots = |x_n|$ и $a_{pk} x_k = |a_{pk} x_k|$ $(k = 1, \ldots, n)$, где p — то значение i, при котором достигается

максимум в правой части соотношения (20). Таким образом, эта правая часть равна $\sup\limits_{x\neq 0}\frac{\|Ax\|}{\|x\|}$ и имеет место формула (20)[1]).

Если же исходить из октаэдрических векторных норм

$$\|x\|_{\text{II}}=\sum_{k=1}^{n}|x_k|,\qquad \|y\|_{\text{II}}=\sum_{i=1}^{m}|y_i|,$$

то, как нетрудно показать,

$$\|A\|=\max_{1\leqslant k\leqslant n}\sum_{i=1}^{m}|a_{ik}|. \tag{20'}$$

Рассмотрим теперь эрмитовы векторные нормы: $\|x\|^2=\sum\limits_{k=1}^{n}|x_k|^2$ и $\|y\|^2=\sum\limits_{i=1}^{n}|y_i|^2$. Тогда, вводя в рассмотрение положительную эрмитову матрицу $S=A^*A$, будем иметь:

$$\|Ax\|^2=y^*y=x^*A^*Ax=x^*Sx,\quad \|x\|^2=x^*x.$$

Но тогда (см. гл. X, § 7)

$$\|A\|^2=\max_{x\neq 0}\frac{x^*Sx}{x^*x}=\varrho,$$

где ϱ — максимальное характеристическое число матрицы AA^*. В этом случае

$$\|A\|=\sqrt{\varrho}. \tag{20''}$$

Введем теперь различные нормы для векторных столбцов x и y. Пусть, например,

$$\|x\|_{\text{II}}=\sum_{k=1}^{n}|x_k|,\qquad \|y\|_{\text{I}}=\max_{1\leqslant i\leqslant n}|y_i|.$$

Тогда

$$\|Ax\|_{\text{I}}=\max_{1\leqslant i\leqslant m}\left|\sum_{k=1}^{n}a_{ik}x_k\right|\leqslant a\sum_{k=1}^{n}|x_k|=a\|x\|_{\text{II}},$$

---

[1]) Иногда норму квадратной $n\times n$-матрицы A вводят аксиоматически (независимо от векторной нормы): каждой $n\times n$-матрице A ставится в соответствие неотрицательное действительное число $\|A\|$ так, что:

1° $\|A\|>0$, если $A\neq 0$, и $\|A\|\neq 0$, если $A=0$.

2° $\|A+B\|\leqslant\|A\|+\|B\|$.

3° $\|\lambda A\|=|\lambda|\,\|A\|$ (λ — скаляр).

4° $\|AB\|\leqslant\|A\|\,\|B\|$.

Например, можно положить

$$\|A\|=\sqrt{\sum_{i,\,k}|a_{ik}|^2}\quad\text{или}\quad \|A\|=n\max_{i,\,j}|a_{ij}|.$$

Вводят одну и ту же векторную норму как для вектора x, так и для вектора $y=Ax$ и называют эту норму согласованной с нормой $\|A\|$, если всегда выполняется соотношение $\|Ax\|\leqslant\|A\|\,\|x\|$. Наше определение нормы в этом частном случае ($m=n$ и $R\equiv S$) удовлетворяет требованиям 1°—4° и согласуется с векторной нормой. В отличие от произвольной нормы, определяемой аксиоматически, норму матрицы, определяемую с помощью формулы (18), называют операторной нормой, подчиненной данной норме векторов.

где $a = \max\limits_{1 \leqslant i,\, k \leqslant n} |a_{ik}|$. С другой стороны, если $a = a_{pq}$, то, выбирая x_q так, чтобы $a_{pq}x_q = a|x_q|$, и полагая $x_j = 0$ при $j \neq q$, будем иметь равенство $\|Ax\|_{\mathrm{I}} = a\|x\|_{\mathrm{II}}$.

Таким образом, в этом случае

$$\|Ax\| = \max_{1 \leqslant i,\, k \leqslant n} |a_{ik}|. \tag{20'''}$$

§ 3. Распространение критерия Адамара на блочные матрицы

Пусть $n \times n$-матрица A разбита на s^2 блоков $A_{\alpha\beta}$ с размерами соответственно $n_\alpha \times n_\beta$ ($\alpha, \beta = 1, \ldots, s$):

$$A = \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & & n_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1s} \\ A_{21} & A_{22} & \ldots & A_{2s} \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{pmatrix} \end{array} \begin{array}{l} \}\, n_1 \\ \}\, n_2 \\ \\ \}\, n_s \end{array}. \tag{21}$$

При этом n-мерное пространство R автоматически расщепляется на s подпространств R_α с числом измерений n_α ($i = 1, \ldots, s$). Для любого вектора $x \in R$ имеет место разложение

$$x = \sum_{\alpha=1}^{s} x_\alpha \qquad (x_\alpha \in R_\alpha, \quad \alpha = 1, \ldots, s). \tag{21'}$$

Введем векторные нормы в пространствах R_α. Поскольку блок-матрица $A_{\alpha\beta}$ отображает R_β в R_α, то тем самым определится и норма

$$\|A_{\alpha\beta}\| = \sup_{\substack{x_\beta \in R_\beta \\ x_\beta \neq 0}} \frac{\|A_{\alpha\beta}x_\beta\|_{R_\alpha}}{\|x_\beta\|_{R_\beta}}. \tag{22}$$

В частности, определяется и норма квадратных матриц $A_{\alpha\alpha}$:

$$\|A_{\alpha\alpha}\| = \sup_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|A_{\alpha\alpha}x_\alpha\|}{\|x_\alpha\|}. \tag{22'}$$

Если $|A_{\alpha\alpha}| \neq 0$, то $\|A_{\alpha\alpha}\| > 0$. В этом случае из (22′) легко следует, что

$$\|A_{\alpha\alpha}^{-1}\| = \sup_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|x_\alpha\|}{\|A_{\alpha\alpha}x_\alpha\|}$$

и, следовательно,

$$\|A_{\alpha\alpha}^{-1}\|^{-1} = \inf_{\substack{x_\alpha \in R_\alpha \\ x_\alpha \neq 0}} \frac{\|A_{\alpha\alpha}x_\alpha\|}{\|x_\alpha\|}. \tag{23}$$

Правая часть этого равенства имеет смысл и в случае, когда $A_{\alpha\alpha}$ — вырожденная матрица. (В этом случае справа стоит нуль.) Исходя из этого и из соображений непрерывности, будем считать, что $\|A_{\alpha\alpha}^{-1}\|^{-1}$ определено и в случае $|A_{\alpha\alpha}| = 0$ равно нулю.

Пусть теперь $|A| = 0$ и имеет место равенство $Ax = 0$ при $x \neq 0$. Исходя из представлений (21) и (21′), раскрывая блочное произведение

Ax, мы сможем написать

$$-A_{\alpha\alpha}x_\alpha = \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} A_{\alpha\beta}x_\beta \qquad (\alpha=1, \ldots, s). \tag{24}$$

Отсюда в силу установленных ранее свойств нормы матрицы (см. (18′) и (19))

$$\| A_{\alpha\alpha}x_\alpha \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta}x_\beta \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| \, \| x_\beta \| \qquad (\alpha=1, \ldots, s). \tag{25}$$

С другой стороны, из (23) следует

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \| x_\alpha \| \leqslant \| A_{\alpha\alpha}x_\alpha \| \qquad (\alpha=1, \ldots, s),$$

что в сочетании с предыдущими неравенствами (25) дает

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \| x_\alpha \| \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| \, \| x_\beta \| \qquad (\alpha=1, \ldots, s). \tag{26}$$

Как и в § 1, выберем индекс α так, чтобы $\| x_\alpha \|$ имел наибольшее значение (по сравнению с $\| x_\beta \|$, где $\beta \neq \alpha$), и заменим в правой части (26) все $\| x_\beta \|$ на $\| x_\alpha \|$. После сокращения на $\| x_\alpha \| > 0$ получим

$$\| A_{\alpha\alpha}^{-1} \|^{-1} \leqslant \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \|. \tag{27}$$

Поэтому при выполнении «блочных условий Адамара»

$$\| A_{\alpha\alpha}^{-1} \|^{-1} - \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A_{\alpha\beta} \| > 0 \qquad (\alpha=1, \ldots, s) \tag{28}$$

соотношение (27) невозможно и матрица A не может быть вырожденной.

Мы пришли к теореме:

Теорема 3. *Если выполняются блочные условия Адамара* (28), *то A — невырожденная матрица.*

В частном случае $n_1 = n_2 = \ldots = n_s = 1$ эта теорема переходит в теорему Адамара, если в одномерных подпространствах R_α определить норму так: $\| x_\alpha \| = | x_\alpha |$ $(\alpha=1, \ldots, s)$.

Само собой разумеется, что, записывая условие невырожденности транспонированной матрицы A', можно в теореме 3 условия Адамара для блочных строк заменить условиями Адамара для блочных столбцов:

$$\| A_{\alpha\alpha}^{-1} \|^{-1} - \sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s} \| A'_{\beta\alpha} \| > 0 \qquad (\alpha=1, \ldots, s). \tag{28′}$$

На блочные матрицы легко распространяется и теорема Ольги Тауски, если только в этой теореме потребовать «блочную» неприводимость матрицы A и выполнение ослабленных блочных условий Адамара со строгим знаком $>$ хотя бы в одном из них.

§ 4. Критерий регулярности Фидлера

Пусть снова $n \times n$-матрица A представлена в блочном виде (21). Составим для нее числовую $s \times s$-матрицу с вещественными элементами:

$$G = \begin{pmatrix} \|A_{11}^{-1}\|^{-1} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \|A_{22}^{-1}\| & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \|A_{ss}^{-1}\|^{-1} \end{pmatrix}. \tag{29}$$

У этой матрицы все недиагональные элементы $\leqslant 0$, а диагональные $\geqslant 0$. Напомним читателю, что матрица с вещественными элементами называется M-матрицей, если у нее все недиагональные элементы $\leqslant 0$, т. е. неположительные и все главные миноры положительны[1]). Имеет место

Теорема 4 (Фидлера). *Если $s \times s$-матрица G является M-матрицей, то $n \times n$-матрица A является регулярной.*

Доказательство. Допустим, что $|A| = 0$. Тогда $Ax = 0$, где $x \neq 0$. Исходя из представлений (21) и (21′), как и ранее на стр. 413, получаем неравенства (26), которые теперь перепишем так:

$$\|A_{\alpha\alpha}^{-1}\|^{-1}\|x_\alpha\| - \sum_{\substack{\beta=1 \\ \beta\neq\alpha}}^{s} \|A_{\alpha\beta}\|\,\|x_\beta\| \leqslant 0 \qquad (\alpha = 1, \dots, s). \tag{30}$$

1. Пусть сначала все $\|x_\alpha\| > 0$. Тогда, увеличивая надлежащим образом в (30) коэффициент при $\|x_\alpha\|$, т. е. заменяя $\|A_{\alpha\alpha}^{-1}\|^{-1}$ на некоторое число $\tilde{g}_{\alpha\alpha} \geqslant \|A_{\alpha\alpha}^{-1}\|^{-1}$, мы из неравенств (30) получим систему равенств

$$\tilde{g}_{\alpha\alpha}\|x_\alpha\| - \sum_{\substack{\beta=1 \\ \beta\neq\alpha}}^{s} \|A_{\alpha\beta}\|\,\|x_\beta\| = 0 \qquad (\alpha = 1, \dots, s),$$

которые в матричной символике запишутся так:

$$\tilde{G}\xi = 0,$$

где

$$\tilde{G} = \begin{pmatrix} \tilde{g}_{11} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \tilde{g}_{22} & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \tilde{g}_{ss} \end{pmatrix},$$

а $\xi \neq 0$ — s-мерный вектор-столбец с элементами $\|x_1\|, \dots, \|x_s\|$. Отсюда сразу следует, что $|\tilde{G}| = 0$. С другой стороны, из определения M-матрицы следует, что $|\tilde{G}| \geqslant |G| > 0$. Мы пришли к противоречию, допустив, что $|A| = 0$.

Если некоторые $\|x_\alpha\| = 0$, то мы возьмем лишь те из соотношений (30), которые соответствуют значениям α, при которых $\|x_\alpha\| > 0$.

---

[1]) Согласно лемме Котелянского (см. стр. 369), для этого достаточно, чтобы n последовательных главных миноров были положительны.

Повторяя дословно предыдущие рассуждения и оперируя вместо $|G|$ некоторым главным минором матрицы G, мы снова придем к противоречию.

Теорема доказана полностью.

§ 5. Круги Гершгорина и другие области локализации

Пусть $A=\|a_{ik}\|_1^n$ — произвольная $n\times n$-матрица с комплексными элементами и λ — некоторое ее характеристическое число. Тогда $A-\lambda E$ — вырожденная матрица, и потому для нее не могут выполняться все условия Адамара, т. е. должно иметь место хотя бы одно из соотношений

$$|a_{ii}-\lambda|\leqslant\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}| \qquad (i=1,\ \ldots,\ n). \tag{31}$$

Каждое из соотношений (31) определяет некоторый круг в комплексной λ-плоскости с центром в точке a_{ii} радиуса $\sum\limits_{j=1,\ j\neq i}^{n}|a_{ij}|$. Мы пришли к теореме, установленной Гершгориным в 1931 г.

Теорема 5 (Гершгорина). *Каждое характеристическое число λ матрицы $A=\|a_{ik}\|_1^n$ всегда расположено в одном из кругов* (31).

Таким образом, объединение всех точек кругов Гершгорина (31) дает некоторую *область локализации* характеристических чисел матрицы A, т. е. область, в которой заведомо лежат все характеристические числа матрицы A. Каждый критерий регулярности приводит к своей области локализации характеристических чисел. Так, исходя из условий Адамара для столбцов, мы получили область локализации в виде объединения n кругов

$$|\lambda-a_{ii}|\leqslant\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}| \qquad (i=1,\ \ldots,\ n). \tag{31'}$$

Блочные условия Адамара приводят нас сразу к теореме:

Теорема 6. *Каждое характеристическое число λ $n\times n$-матрицы A, представленной в блочном виде $A=\{A_{\alpha\beta}\}_1^s$, принадлежит по крайней мере одной из областей:*

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}\leqslant\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s}\|A_{\alpha\beta}\| \qquad (\alpha=1,\ \ldots,\ s), \tag{32}$$

а также по крайней мере одной из областей

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}\leqslant\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{s}\|A_{\beta\alpha}\| \qquad (\alpha=1,\ \ldots,\ s) \tag{32'}$$

(здесь E_α — единичная матрица того же порядка, что и $A_{\alpha\alpha}$; $\alpha=1,\ \ldots,\ s$).

Выясним, какую область локализации можно получить, исходя из критерия Фидлера. Пусть неотрицательные числа $c_1,\ c_2,\ \ldots,\ c_s$

выбраны так, чтобы матрица

$$G=\begin{pmatrix} c_1 & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & c_2 & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & c_s \end{pmatrix} \tag{33}$$

была *ослабленной M-матрицей*, т. е. чтобы все главные миноры этой матрицы были неотрицательны (недиагональные элементы в этой матрице заведомо все $\leqslant 0$). Допустим теперь, что при некотором числе λ выполняются s неравенств:

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1} > c_\alpha \qquad (\alpha=1, \dots, s). \tag{34}$$

Тогда, заменяя в матрице (33) c_α на $\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1}$ $(\alpha=1, \dots, s)$, мы строго увеличим все диагональные элементы и получим уже M-матрицу (неослабленную!)

$$\begin{pmatrix} \|(A_{11}-\lambda E_1)^{-1}\|^{-1} & -\|A_{12}\| & \dots & -\|A_{1s}\| \\ -\|A_{21}\| & \|(A_{22}-\lambda E_2)^{-1}\|^{-1} & \dots & -\|A_{2s}\| \\ \dots & \dots & \dots & \dots \\ -\|A_{s1}\| & -\|A_{s2}\| & \dots & \|(A_{ss}-\lambda E_s)^{-1}\|^{-1} \end{pmatrix}.$$

Но тогда по теореме Фидлера $|A-\lambda E|\neq 0$ и число λ не является характеристическим числом матрицы A.

Поэтому для любого характеристического числа λ матрицы A по крайней мере одно из неравенств (34) не выполняется, т. е. выполняется одно из соотношений

$$\|(A_{\alpha\alpha}-\lambda E_\alpha)^{-1}\|^{-1} \leqslant c_\alpha \qquad (\alpha=1, \dots, s). \tag{35}$$

Объединение s областей (35) и образует область локализации Фидлера, зависящую от специально выбираемых неотрицательных параметров $c_1, c_2, \dots, c_s$.

Теорема 7 (Фидлера). *Если неотрицательные числа $c_1, c_2, \dots, c_s$ выбраны так, чтобы матрица* (33) *была ослабленной M-матрицей, то каждое характеристическое число λ матрицы A принадлежит по крайней мере одной из s замкнутых областей* (35).

Рассмотрим в качестве примера симметрическую матрицу 4-го порядка

$$A=\begin{Vmatrix} 0 & 4 & 1 & -1 \\ 4 & 0 & -1 & 1 \\ 1 & -1 & -1 & 15 \\ -1 & 1 & 15 & -1 \end{Vmatrix}.$$

Поскольку матрица A симметрическая, то у нее все характеристические числа вещественны. Поэтому вместо областей локализации в комплексной λ-плоскости можно рассматривать отрезки, высекаемые этими областями на вещественной λ-оси.

I. Область Гершгорина состоит из одного сегмента

$$-18 \leqslant \lambda \leqslant 16,$$

который перекрывают остальные сегменты Гершгорина.

II. Разобьем матрицу A на 4 блока:

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad A_{11}=\begin{Vmatrix} 0 & 4 \\ 4 & 0 \end{Vmatrix}, \quad A_{22}=\begin{Vmatrix} -1 & 15 \\ 15 & -1 \end{Vmatrix}, \quad A_{12}=A_{21}=\begin{Vmatrix} 1 & -1 \\ -1 & 1 \end{Vmatrix}.$$

В данном случае

$$(A_{11}-\lambda E_1)=\frac{1}{\lambda^2-16}\begin{Vmatrix}-\lambda & -4\\ -4 & -\lambda\end{Vmatrix},$$

$$(A_{22}-\lambda E_2)^{-1}=\frac{1}{(\lambda+1)^2-15^2}\begin{Vmatrix}-1-\lambda & -1-\lambda\\ -15 & -15\end{Vmatrix}.$$

Рассмотрим три варианта нормировки подпространств R_1 и R_2:

а) в R_1 и R_2 — кубические нормы;

б) в R_1 кубическая, а в R_2 — октаэдральная норма;

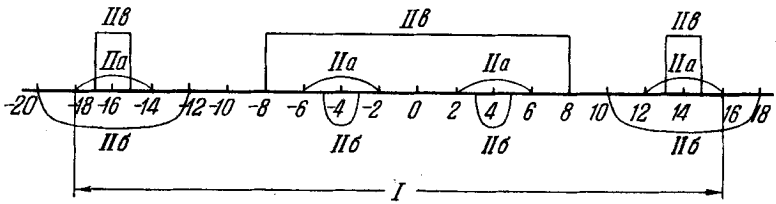

Рис. 6.

в) в R_1 октаэдральная, а в R_2 — кубическая норма.

а) Нормы всех блоков определяются по формуле (20′):

$$\|A_{12}\|=2,\ \|A_{21}\|=2,\ \|(A_{11}-\lambda E_1)^{-1}\|=||\lambda|-4|,\ \|(A_{22}-\lambda E_2)^{-1}\|^{-1}=||\lambda+1|-15|.$$

Блочные области Гершгорина:

$$||\lambda|-4|\leqslant 2,\qquad ||\lambda+1|-15|\leqslant 2$$

представляют собой совокупность 4 интервалов:

$$-18\leqslant\lambda\leqslant-16,\quad -6\leqslant\lambda\leqslant-2,\quad 2\leqslant\lambda\leqslant 6,\quad 12\leqslant\lambda\leqslant 16.\qquad \text{(IIа)}$$

б) В этом случае выражения для $\|(A_{11}-\lambda E_1)^{-1}\|^{-1}$ и $\|(A_{22}-\lambda E_2)^{-1}\|^{-1}$ остаются прежними, но

$$\|A_{12}\|=\max_x\frac{|x_1-x_2|}{|x_1|+|x_2|}=1,\quad \|A_{21}\|=\max_x\frac{2|x_1-x_2|}{\max\limits_{1\leqslant i\leqslant 2}|x_i|}=4.$$

Блочные области Гершгорина:

$$||\lambda|-4|\leqslant 1,\quad ||\lambda+1|-15|\leqslant 4$$

распадаются на 4 интервала

$$-20\leqslant\lambda\leqslant-12,\quad -5\leqslant\lambda\leqslant-3,\quad 3\leqslant\lambda\leqslant 5,\quad 10\leqslant\lambda\leqslant 18.\qquad \text{(IIб)}$$

в) Отличие от предыдущего случая заключается лишь в том, что здесь

$$\|A_{21}\|=4,\quad \|A_{12}\|=1.$$

Поэтому блочные области Гершгорина

$$||\lambda|-4|\leqslant 4,\quad ||\lambda+1|-15|\leqslant 1$$

распадаются на 3 интервала:

$$-17\leqslant\lambda\leqslant-15,\quad -8\leqslant\lambda\leqslant 8,\quad 13\leqslant\lambda\leqslant 15.\qquad \text{(IIв)}$$

На схеме (рис. 6) изображены области I, IIа, IIб, IIв. Их пересечение дает области локализации:

$$-17\leqslant\lambda\leqslant-15,\quad -5\leqslant\lambda\leqslant-3,\quad 3\leqslant\lambda\leqslant 5,\quad 13\leqslant\lambda\leqslant 15.$$

III. При применении критерия Фидлера будем снова исходить из нормировок а), б), в):

$$G=\begin{pmatrix} c_1 & -\|A_{12}\| \\ -\|A_{21}\| & c_2 \end{pmatrix}=\begin{Vmatrix} c_1 & -2 \\ -2 & c_2 \end{Vmatrix}, \quad |G|=c_1c_2-4\geqslant 0$$

Желая получить наименьшие значения c_1 и c_2, полагаем

$$c_1c_2=4.$$

Область Фидлера

$$||\lambda|-4|\leqslant c_1, \quad ||\lambda+1|-15|\leqslant c_2$$

совпадает с областью IIа при $c_1=2$, $c_2=2$, с областью IIб — при $c_1=1$,

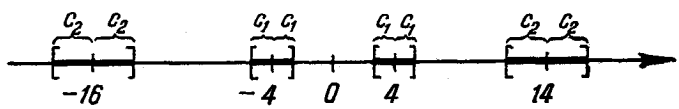

Рис. 7.

$c_2=4$ и с областью IIв при $c_1=4$, $c_2=1$. Область Фидлера состоит из 4 интервалов:

$$\left.\begin{array}{ll} -16-c_2\leqslant\lambda\leqslant -16+c_2, & -4-c_1\leqslant\lambda\leqslant -4+c_1, \\ 4-c_1\leqslant\lambda\leqslant 4+c_1, & 14-c_2\leqslant\lambda\leqslant 14+c_2 \end{array}\right\} \qquad \text{(III)}$$

и зависит от одного положительного параметра, поскольку $c_1=4/c_2$. Можно определить пересечение всех этих областей Фидлера.

Для этого (рис. 7) приравняем величины:

1) $-16-c_2=-4-c_1$, 2) $-16+c_2=-4-c_1$,

3) $-16+c_2=-4+c_1$, 4) $4-c_1=14-c_2$,

5) $4+c_1=14-c_2$, 6) $4+c_1=14+c_2$.

Используя равенство $c_1c_2=4$, получим 6 квадратных уравнений с наименьшими положительными корнями:

1) $c_2^2-12c_2-4=0$, $z_1=-6+\sqrt{40}=0{,}3246\ldots$,

2) $c_2^2-12c_2+4=0$, $z_2=6-\sqrt{32}=0{,}3431\ldots$,

3) $c_1^2+12c_1-4=0$, $z_3=z_1=-6+\sqrt{40}=0{,}3246\ldots$,

4) $c_1^2+10c_1-4=0$, $z_4=-5+\sqrt{29}=0{,}3852\ldots$,

5) $c_1^2-10c_1+4=0$, $z_5=5-\sqrt{21}=0{,}4174\ldots$,

6) $c_2^2+10c_2-4=0$, $z_6=z_4=-5+\sqrt{29}=0{,}3852\ldots$

Нетрудно уяснить себе, что область локализации, состоящая из пересечения всех областей Фидлера, состоит из следующих 4 сегментов:

$$-16-z_1\leqslant\lambda\leqslant -16+z_2, \qquad -4-z_2\leqslant\lambda\leqslant -4+z_1,$$
$$4-z_4\leqslant\lambda\leqslant 4+z_5, \qquad 14-z_5\leqslant\lambda\leqslant 14+z_4.$$

ГЛАВА XV

ПРИЛОЖЕНИЯ ТЕОРИИ МАТРИЦ К ИССЛЕДОВАНИЮ СИСТЕМ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

§ 1. Системы линейных дифференциальных уравнений с переменными коэффициентами. Общие понятия

Пусть дана система линейных однородных дифференциальных уравнений первого порядка:

$$\frac{dx_i}{dt}=\sum_{k=1}^{n} p_{ik}(t)\,x_k \qquad (i=1,\ 2,\ \ldots,\ n), \tag{1}$$

где $p_{ik}(t)$ $(i,\ k=1,\ 2,\ \ldots,\ n)$ — комплексные функции вещественного аргумента t, непрерывные в некотором (конечном или бесконечном) интервале изменения t [1]).

Полагая $P(t)=\|p_{ik}(t)\|_1^n$ и $x=(x_1,\ x_2,\ \ldots,\ x_n)$, мы систему (1) запишем так:

$$\frac{dx}{dt}=P(t)\,x. \tag{2}$$

Интегральной матрицей системы (1) мы будем называть квадратную матрицу $X(t)=\|x_{ik}(t)\|_1^n$, столбцами которой являются n линейно независимых решений системы.

Так как каждый столбец матрицы X удовлетворяет уравнению (2), то и интегральная матрица X удовлетворяет уравнению

$$\frac{dX}{dt}=P(t)\,X. \tag{3}$$

В дальнейшем мы вместо системы (1) будем рассматривать матричное уравнение (3).

Из теоремы о существовании и единственности решения системы линейных дифференциальных уравнений [2]) следует, что интегральная матрица $X(t)$ однозначно определяется, если задано значение этой матрицы при некотором («начальном») значении $t=t_0$ [3]), $X(t_0)=X_0$. В качестве матрицы X_0 можно взять любую неособенную квадратную матрицу n-го порядка. В частном случае, когда $X(t_0)=E$, интегральную матрицу $X(t)$ будем называть *нормированной*.

---

[1]) Все соотношения этого параграфа, в которые входят функции от t, имеют место для данного интервала изменения t.

[2]) Доказательство этой теоремы приведено далее в § 5. См. также И. Г Петровский, Лекции по теории обыкновенных дифференциальных уравнений, «Наука», 1964.

[3]) Предполагается, что t_0 принадлежит данному интервалу изменения t.

Продифференцируем определитель матрицы X, дифференцируя последовательно строки определителя и используя при этом дифференциальные соотношения

$$\frac{dx_{ij}}{dt}=\sum_{k=1}^{n} p_{ik}x_{kj} \qquad (i,\ j=1,\ 2,\ \ldots,\ n).$$

Тогда получим:

$$\frac{d|X|}{dt}=(p_{11}+p_{22}+\ldots+p_{nn})\,|X|.$$

Отсюда следует известное тождество Якоби

$$|X|=ce^{\int\limits_{t_0}^{t}\operatorname{Sp}P\,dt}, \tag{4}$$

где c — постоянная, а

$$\operatorname{Sp}P=p_{11}+p_{22}+\ldots+p_{nn}$$

— след матрицы $P(t)$.

Так как определитель $|X|$ не может тождественно равняться нулю, то $c\neq 0$. Но тогда из тождества Якоби следует, что определитель $|X|$ при любом значении аргумента отличен от нуля

$$|X|\neq 0,$$

т. е. *интегральная матрица при любом значении аргумента является неособенной.*

Если $\widetilde{X}(t)$ — неособенное ($|\widetilde{X}|\neq 0$) частное решение уравнения (3), то общее решение этого уравнения определяется формулой

$$X=\widetilde{X}C, \tag{5}$$

где C — произвольная постоянная матрица.

Действительно, умножая обе части равенства

$$\frac{d\widetilde{X}}{dt}=P\widetilde{X} \tag{6}$$

справа на C, убеждаемся, что и матрица $\widetilde{X}C$ удовлетворяет уравнению (3). С другой стороны, если X — произвольное решение уравнения (3), то из (6) следует:

$$\frac{dX}{dt}=\frac{d}{dt}(\widetilde{X}\cdot\widetilde{X}^{-1}X)=\frac{d\widetilde{X}}{dt}\widetilde{X}^{-1}X+\widetilde{X}\frac{d}{dt}(\widetilde{X}^{-1}X)=PX+\widetilde{X}\frac{d}{dt}(\widetilde{X}^{-1}X),$$

откуда в силу (3)

$$\frac{d}{dt}(\widetilde{X}^{-1}X)=0$$

и

$$\widetilde{X}^{-1}X=\text{const}=C,$$

т. е. имеет место (5).

Все интегральные матрицы X системы (1) получаются по формуле (5) при $|C|\neq 0$.

Рассмотрим частный случай:

$$\frac{dX}{dt}=AX, \tag{7}$$

где A — постоянная матрица. При этом $\widetilde{X}=e^{At}$ есть частное неособенное решение уравнения (7)[1]) и потому общее решение этого уравнения имеет вид

$$X=e^{At}C, \tag{8}$$

где C — произвольная постоянная матрица.

Полагая в (8) $t=t_0$, найдем: $X_0=e^{At_0}C$. Отсюда $C=e^{-At_0}X_0$ и потому формулу (8) можно представить в виде

$$X=e^{A(t-t_0)}X_0. \tag{9}$$

Эта формула эквивалентна выведенной ранее формуле (46) главы V (стр. 125).

Рассмотрим еще так называемую систему Коши:

$$\frac{dX}{dt}=\frac{A}{t-a}X \qquad (A \text{ — постоянная матрица}). \tag{10}$$

Этот случай сводится к предыдущему заменой аргумента:

$$u=\ln(t-a).$$

Поэтому общее решение системы (10) выглядит так:

$$X=e^{A\ln(t-a)}C=(t-a)^A C. \tag{11}$$

Функции e^{At} и $(t-a)^A$, встречающиеся в формулах (8) и (11), могут быть представлены в виде (стр. 125)

$$e^{At}=\sum_{k=1}^{s}(Z_{k1}+Z_{k2}t+\ldots+Z_{km_k}t^{m_k-1})e^{\lambda_k t}, \tag{12}$$

$$(t-a)^A=\sum_{k=1}^{s}(Z_{k1}+Z_{k2}\ln(t-a)+\ldots+Z_{km_k}[\ln(t-a)]^{m_k-1})(t-a)^{\lambda_k}. \tag{13}$$

Здесь

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_s)^{m_s}$$
$$(\lambda_i\neq\lambda_k \text{ при } i\neq k;\ i,\ k=1,\ 2,\ \ldots,\ s)$$

— минимальный многочлен матрицы A, а Z_{kj} $(j=1,\ 2,\ \ldots,\ m_k;\ k=1,\ 2,\ \ldots,\ s)$ — линейно независимые постоянные матрицы, являющиеся многочленами от A[2]).

Замечание. Иногда в качестве интегральной матрицы системы дифференциальных уравнений (1) берут матрицу W, у которой строки являются линейно независимыми решениями системы. Очевидно, матрица W будет транспонированной матрицей для X:

$$W=X'.$$

---

[1]) Почленно дифференцируя ряд $e^{At}=\sum_{k=0}^{\infty}\frac{A^k}{k!}t^k$, находим: $\frac{d}{dt}e^{At}=Ae^{At}$.

[2]) В правой части формулы (12) каждое слагаемое $X_k=(Z_{k1}+Z_{k2}t+\ldots+Z_{km_k}t^{m_k-1})e^{\lambda_k t}$ $(k=1,\ 2,\ \ldots,\ s)$ является решением уравнения (7). Действительно, этому уравнению удовлетворяет произведение $g(A)e^{At}$ при любой функции $g(\lambda)$. Но $X_k=f(A)=g(A)e^{At}$, если $f(\lambda)=g(\lambda)e^{\lambda t}$ и $g(\lambda_k)=1$, а все остальные $m-1$ значений функции $g(\lambda)$ на спектре матрицы A равны нулю [см. (16) на стр. 111].

Переходя в обеих частях равенства (3) к транспонированным матрицам, мы вместо (3) получим следующее уравнение для W:

$$\frac{dW}{dt} = W P(t). \tag{3'}$$

В правой части этого уравнения матрица W стоит первым множителем, а не вторым, как X в уравнении (3).

§ 2. Преобразование Ляпунова

Допустим теперь, что в системе (1) [и в уравнении (3)] матрица коэффициентов $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная ограниченная функция от t в интервале $[t_0, \infty)$ [1]).

Введем вместо неизвестных функций $x_1, x_2, \ldots, x_n$ новые неизвестные функции $y_1, y_2, \ldots, y_n$ при помощи преобразования

$$x_i = \sum_{k=1}^{n} l_{ik}(t) y_k \qquad (i = 1, 2, \ldots, n). \tag{14}$$

На матрицу преобразования $L(t) = \| l_{ik}(t) \|_1^n$ наложим следующие ограничения:

1° $L(t)$ имеет непрерывную производную $\frac{dL}{dt}$ в интервале $[t_0, \infty)$;

2° $L(t)$ и $\frac{dL}{dt}$ ограничены в интервале $[t_0, \infty)$;

3° существует постоянная m такая, что

$$0 < m < \operatorname{mod} |L(t)| \qquad (t \geqslant t_0),$$

т. е. определитель $|L(t)|$ ограничен по модулю снизу положительной постоянной m.

Преобразование (14), в котором матрица коэффициентов $L(t) = \| l_{ik}(t) \|_1^n$ удовлетворяет условиям 1°—3°, мы будем называть *преобразованием Ляпунова*, а соответствующую матрицу $L(t)$ — *матрицей Ляпунова*.

Такие преобразования рассматривал А. М. Ляпунов в своем знаменитом мемуаре «Общая задача об устойчивости движения» [19].

Примеры. 1. Если $L = \text{const}$ и $|L| \neq 0$, то матрица L удовлетворяет условиям 1°—3°. Следовательно, неособенное преобразование с постоянными коэффициентами всегда является преобразованием Ляпунова.

2. Если $D = \| d_{ik} \|_1^n$ — матрица простой структуры с чисто мнимыми характеристическими числами, то матрица

$$L(t) = e^{Dt}$$

удовлетворяет условиям 1°—3° и потому является матрицей Ляпунова [2]).

Легко проверить, что из свойств 1°—3° матрицы $L(t)$ следует, что существует обратная матрица $L^{-1}(t)$ и что она удовлетворяет тем же условиям 1°—3°, т. е. обратное преобразование для преобразования Ляпунова снова является преобразованием Ляпунова. Точно так же проверяется, что два последовательных преобразования Ляпунова в результате снова дают преобразование Ляпунова. Таким образом, преобразования Ляпунова образуют группу. Преобразования Ляпунова обладают следующим важным свойством:

---

[1]) Это означает, что каждая из функций $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ непрерывна и ограничена в интервале $[t_0, \infty)$, т. е. при $t \geqslant t_0$.

[2]) При этом в формуле (12) все $m_k = 1$, а $\lambda_k = i\varphi_k$ (φ_k вещественны, $k = 1, 2, \ldots, s$).

Если при преобразовании (14) *система уравнений* (1) *переходит в систему*

$$\frac{dy_i}{dt} = \sum_{k=1}^{n} q_{ik}(t)\, y_k, \tag{15}$$

нулевое решение которой является устойчивым, асимптотически устойчивым или неустойчивым по Ляпунову (см. гл. V, § 7), *то таким же свойством обладает и нулевое решение исходной системы* (1).

Другими словами, преобразования Ляпунова не изменяют характеристики нулевого решения (в отношении устойчивости). Поэтому эти преобразования могут быть использованы при исследовании устойчивости для упрощения исходной системы уравнений.

Преобразование Ляпунова устанавливает одно-однозначное соответствие между решениями систем (1) и (15), при этом линейно независимые решения остаются таковыми и после преобразования. Поэтому преобразование Ляпунова переводит интегральную матрицу X системы (1) в некоторую интегральную матрицу Y системы (15), при этом

$$X = L(t)\, Y. \tag{16}$$

В матричной записи система (15) имеет вид

$$\frac{dY}{dt} = Q(t)\, Y, \tag{17}$$

где $Q(t) = \| q_{ik}(t) \|_1^n$ — матрица коэффициентов системы (15).

Подставляя в (3) вместо X произведение LY и сопоставляя полученное уравнение с (17), легко найдем следующую формулу, выражающую матрицу Q через матрицы P и L:

$$Q = L^{-1}PL - L^{-1}\frac{dL}{dt}. \tag{18}$$

Две системы (1) и (15) или, что то же, (3) и (17) мы будем называть *эквивалентными* (в смысле Ляпунова), если они переводятся друг в друга преобразованием Ляпунова. Матрицы коэффициентов P и Q эквивалентных систем всегда связаны между собой формулой (18), в которой матрица L удовлетворяет условиям 1°—3°.

§ 3. Приводимые системы

Среди систем линейных дифференциальных уравнений первого порядка наиболее простыми и наиболее изученными являются системы с постоянными коэффициентами. Поэтому представляют интерес системы, которые при помощи преобразования Ляпунова могут быть приведены к системам с постоянными коэффициентами. Такие системы А. М. Ляпунов называл *приводимыми*.

Пусть дана приводимая система

$$\frac{dX}{dt} = PX. \tag{19}$$

Тогда некоторое преобразование Ляпунова

$$X = L(t)\, Y \tag{20}$$

переводит ее в систему

$$\frac{dY}{dt} = AY, \tag{21}$$

где A — постоянная матрица. Поэтому система (19) имеет частное решение

$$\widetilde{X} = L(t)\, e^{At}. \tag{22}$$

Легко видеть, что, и обратно, всякая система (19), имеющая частное решение вида (22), где $L(t)$ — матрица Ляпунова, а A — постоянная матрица, является приводимой и при этом она приводится к виду (21) при помощи преобразования Ляпунова (20).

Следуя А. М. Ляпунову, покажем, что *всякая система* (19) *с периодическими коэффициентами приводима* [1]).

Пусть в данной системе (19) $P(t)$ — непрерывная функция в интервале $(-\infty, +\infty)$ с периодом τ:

$$P(t+\tau) = P(t). \tag{23}$$

Заменяя в (19) t на $t+\tau$ и используя (23), получим:

$$\frac{dX(t+\tau)}{dt} = P(t)\, X(t+\tau).$$

Таким образом, $X(t+\tau)$, как и $X(t)$, является интегральной матрицей системы (19). Поэтому $X(t+\tau) = X(t)\, V$, где V — некоторая постоянная неособенная матрица. Поскольку $|V| \neq 0$, то можно определить [2])

$$V^{\frac{t}{\tau}} = e^{\frac{t}{\tau}\ln V}.$$

Эта матричная функция от t, как и $X(t)$, умножается справа на V, если к аргументу прибавить τ. Поэтому «частное»

$$L(t) = X(t)\, V^{-\frac{t}{\tau}} = X(t)\, e^{-\frac{t}{\tau}\ln V}$$

является непрерывной периодической функцией с периодом τ:

$$L(t+\tau) = L(t)$$

и с определителем $|L(t)| \neq 0$. Матрица $L(t)$ удовлетворяет условиям 1°—3° предыдущего параграфа и, следовательно, является матрицей Ляпунова.

С другой стороны, поскольку решение X системы (19) представимо в виде

$$X = L(t)\, e^{\frac{\ln V}{\tau} t},$$

то система (19) приводима.

В данном случае преобразование Ляпунова

$$X = L(t)\, Y,$$

приводящее систему (19) к виду

$$\frac{dY}{dt} = \frac{1}{\tau}\ln V \cdot Y,$$

имеет периодические коэффициенты с периодом τ.

А. М. Ляпуновым был установлен [3]) весьма важный критерий устойчивости и неустойчивости по первому линейному приближению для

---

[1]) См. [22], § 47.

[2]) Здесь $\ln V = f(V)$, где $f(\lambda)$ — какая-либо однозначная ветвь функции $\ln \lambda$ в односвязной области G, содержащей все характеристические числа матрицы V и не содержащей числа 0. См. гл. V.

[3]) См. [22], § 24.

нелинейных систем дифференциальных уравнений

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} a_{ik}x_k + (**) \qquad (i = 1, 2, \ldots, n), \tag{24}$$

где в правых частях стоят сходящиеся степенные ряды относительно $x_1, x_2, \ldots, x_n$, а $(**)$ обозначает сумму членов этих рядов второго порядка и выше относительно $x_1, x_2, \ldots, x_n$; коэффициенты a_{ik} $(i, k = 1, 2, \ldots, n)$ в линейных членах постоянны [1]).

Критерий Ляпунова. *Нулевое решение системы* (24) *будет устойчивым* (*и притом асимптотически*), *если матрица коэффициентов первого линейного приближения* $A = \| a_{ik} \|_1^n$ *имеет все характеристические числа с отрицательными вещественными частями, и неустойчивым, если хотя бы одно из этих характеристических чисел имеет положительную вещественную часть.*

Приведенные выше рассуждения позволяют использовать этот критерий для систем с периодическими коэффициентами в линейных членах:

$$\frac{dx_i}{dt} = \sum_{k=1}^{n} p_{ik}(t)\, x_k + (**) \qquad (i = 1, \ldots, n). \tag{25}$$

Действительно, на основании предыдущих рассуждений можно при помощи преобразования Ляпунова систему (25) привести к виду (24), где

$$A = \| a_{ik} \|_1^n = \frac{1}{\tau} \ln V,$$

а V — постоянная матрица, на которую умножается интегральная матрица соответствующей линейной системы (19) при сдвиге аргумента на τ. Не нарушая общности, можем считать $\tau > 0$. В силу свойств преобразования Ляпунова нулевое решение исходной системы и нулевое решение преобразованной одновременно являются устойчивыми, асимптотически устойчивыми или неустойчивыми. Но характеристические числа λ_i и ν_i $(i = 1, 2, \ldots, n)$ матриц A и V связаны между собой формулой

$$\lambda_i = \frac{1}{\tau} \ln \nu_i \qquad (i = 1, 2, \ldots, n).$$

Поэтому, применяя критерий Ляпунова к приведенной системе, найдем [2]):

Нулевое решение системы (25) *будет асимптотически устойчивым, если все характеристические числа* $\nu_1, \nu_2, \ldots, \nu_n$ *матрицы* V *по модулю* <1, *и неустойчивым, если хотя бы одно из этих чисел по модулю* >1.

А. М. Ляпунов установил свой критерий устойчивости по линейному приближению для значительно более широкого класса систем, а именно для систем вида (24), у которых система линейного приближения не обязательно система с постоянными коэффициентами, но принадлежит к классу систем, названных Ляпуновым правильными [3]).

Класс правильных линейных систем содержит в себе как часть все приводимые системы.

Критерий неустойчивости для случая, когда первое линейное приближение является правильной системой, был установлен Н. Г. Четаевым [4]).

---

[1]) Коэффициенты при нелинейных членах могут зависеть от t. На эти функциональные коэффициенты налагаются известные ограничения (см. [22], § 11).

[2]) См. там же, § 55.

[3]) Там же, § 9.

[4]) См. [38], стр. 181.

§ 4. Каноническая форма приводимой системы. Теорема Еругина

Пусть даны приводимая система (19) и эквивалентная ей (в смысле Ляпунова) система

$$\frac{dY}{dt}=AY,$$

где A — постоянная матрица.

Нас будет интересовать вопрос, *в какой степени матрица A определяется данной системой* (19). Этот вопрос можно еще сформулировать так:

В каком случае две системы

$$\frac{dY}{dt}=AY \text{ и } \frac{dZ}{dt}=BZ,$$

где A и B — постоянные матрицы, являются эквивалентными по Ляпунову, т. е. переводятся друг в друга преобразованием Ляпунова?

Для того чтобы ответить на этот вопрос, введем понятие о матрицах, имеющих одну и ту же вещественную часть спектра.

Мы будем говорить, что две матрицы A и B n-го порядка *имеют одну и ту же вещественную часть спектра* в том и только в том случае, когда элементарные делители матриц A и B имеют соответственно вид

$$(\lambda-\lambda_1)^{m_1},\ (\lambda-\lambda_2)^{m_2},\ \ldots,\ (\lambda-\lambda_s)^{m_s} \text{ и } (\lambda-\mu_1)^{m_1}, (\lambda-\mu_2)^{m_2},\ \ldots,\ (\lambda-\mu_s)^{m_s},$$

где

$$\operatorname{Re}\lambda_k=\operatorname{Re}\mu_k \quad (k=1,\ 2,\ \ldots,\ s).$$

Имеет место следующая теорема, установленная Н. П. Еругиным[1]:

Теорема 1 (Еругина). *Две системы*

$$\frac{dY}{dt}=AY \text{ и } \frac{dZ}{dt}=BZ \tag{26}$$

(A и B — постоянные матрицы n-го порядка) эквивалентны в смысле Ляпунова в том и только в том случае, если матрицы A и B имеют одну и ту же вещественную часть спектра.

Доказательство. Пусть даны системы (26). Приведем матрицу A к нормальной жордановой форме[2]) (см. гл. VI, § 7)

$$A=T\{\lambda_1E_1+H_1,\ \lambda_2E_2+H_2,\ \ldots,\ \lambda_sE_s+H_s\}T^{-1}, \tag{27}$$

где

$$\lambda_k=\alpha_k+i\beta_k \quad (\alpha_k,\ \beta_k \text{ — вещественные числа; } k=1,\ 2,\ \ldots,\ s). \tag{28}$$

В соответствии с (27) и (28) положим:

$$\left.\begin{aligned} A_1&=T\{\alpha_1E_1+H_1,\ \alpha_2E_2+H_2,\ \ldots,\ \alpha_sE_s+H_s\}T^{-1},\\ A_2&=T\{i\beta_1E_1,\ i\beta_2E_2,\ \ldots,\ i\beta_sE_s\}T^{-1}.\end{aligned}\right\} \tag{29}$$

---

[1]) См. [11], стр. 9—15. Приведенное здесь доказательство теоремы отличается от доказательства Н. П. Еругина.

[2]) E_k — единичная матрица, в H_k элементы первой наддиагонали равны единице, а остальные элементы равны нулю; порядок E_k, H_k равен степени k-го элементарного делителя матрицы A, т. е. m_k $(k=1,\ 2,\ \ldots,\ s)$.

Тогда

$$A = A_1 + A_2, \quad A_1 A_2 = A_2 A_1. \tag{30}$$

Определим матрицу $L(t)$ равенством $A(t) = e^{A_2 t}$. $L(t)$ — матрица Ляпунова (см. пример 2 на стр. 422).

Но частное решение первой из систем (26) в силу (30) имеет вид

$$e^{At} = e^{A_2 t} e^{A_1 t} = L(t)\, e^{A_1 (t)}.$$

Отсюда следует, что первая из систем (26) эквивалентна системе

$$\frac{dU}{dt} = A_1 U, \tag{31}$$

где согласно (29) матрица A_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы A.

Аналогично вторую из систем (26) заменим эквивалентной

$$\frac{dV}{dt} = B_1 V, \tag{32}$$

где матрица B_1 имеет вещественные характеристические числа и ее спектр совпадает с вещественной частью спектра матрицы B.

Наша теорема будет доказана, если мы покажем, что *две системы* (31) *и* (32), *в которых матрицы* A_1 *и* B_1 — *постоянные матрицы с вещественными характеристическими числами, могут быть эквивалентны лишь в том случае, когда матрицы* A_1 *и* B_1 *подобны*[1]).

Пусть преобразование Ляпунова $U = L_1 V$ переводит (31) в (32). Тогда матрица L_1 удовлетворяет уравнению

$$\frac{dL_1}{dt} = A_1 L_1 - L_1 B_1. \tag{33}$$

Это матричное уравнение относительно L_1 эквивалентно системе n^2 дифференциальных уравнений относительно n^2 элементов матрицы L_1. Правая часть в (33) представляет собой линейную операцию над «вектором» L_1 в пространстве n^2 измерений

$$\frac{dL_1}{dt} = \widehat{F}(L_1) \quad [\widehat{F}(L_1) = A_1 L_1 - L_1 B_1]. \tag{33'}$$

Любое характеристическое число линейного оператора $\widehat{F}$ (и соответствующей ему матрицы порядка n^2) представляется в виде разности $\gamma - \delta$, где γ — характеристическое число матрицы A_1, а δ — характеристическое число матрицы B_1[2]). Отсюда следует, что оператор $\widehat{F}$ имеет только вещественные характеристические числа.

---

[1]) Из этого положения следует теорема 1, поскольку эквивалентность систем (31) и (32) означает эквивалентность систем (26), а подобие матриц A_1 и B_1 означает, что эти матрицы имеют одинаковые элементарные делители и потому матрицы A и B имеют одну и ту же вещественную часть спектра.

[2]) В самом деле, пусть Λ_0 — какое-либо характеристическое число оператора $\widehat{F}$. Тогда существует матрица $L \neq 0$ такая, что $\widehat{F}(L) = \Lambda_0 L$ или

$$(A_1 - \Lambda_0 E)\, L = L B_1. \tag{*}$$

Матрицы $A_1 - \Lambda_0 E$ и B_1 имеют хотя бы одно общее характеристическое число, так как в противном случае существовал бы такой многочлен $g(\lambda)$, что

$$g(A_1 - \Lambda_0 E) = 0, \quad g(B_1) = E,$$

а это невозможно, поскольку из (*) следует: $g(A_1 - \Lambda_0 E) \cdot L = L \cdot g(B_1)$ и $L \neq 0$. Но

Обозначим через $\widehat{\psi}(\lambda)=(\lambda-\widehat{\lambda}_1)^{\widehat{m}_1}(\lambda-\widehat{\lambda}_2)^{\widehat{m}_2}\ldots(\lambda-\widehat{\lambda}_u)^{\widehat{m}_u}$ ($\widehat{\lambda}_i$ вещественны; $\widehat{\lambda}_i\neq\widehat{\lambda}_j$ при $i\neq j$; $i,\ j=1,\ 2,\ \ldots,\ u$) минимальный многочлен для $\widehat{F}$. Тогда решение $L_1(t)=e^{\widehat{F}t}L^{(0)}$ системы (33′) в силу формулы (12) (стр. 421) запишется так:

$$L_1(t)=\sum_{k=1}^{u}\sum_{j=0}^{\widehat{m}_k-1}L_{kj}t^je^{\widehat{\lambda}_kt}, \tag{34}$$

где L_{kj} — постоянные матрицы n-го порядка. Поскольку матрица $L_1(t)$ ограничена в интервале $(t_0,\ \infty)$, то как для любого $\widehat{\lambda}_k>0$, так и при $\widehat{\lambda}_k=0$ и $j>0$ соответствующие матрицы $L_{kj}=0$. Обозначим через $L_-(t)$ сумму всех слагаемых в (34), в которых $\widehat{\lambda}_k<0$. Тогда

$$L_1(t)=L_-(t)+L_0, \tag{35}$$

где

$$\lim_{t\to+\infty}L_-(t)=0,\qquad \lim_{t\to+\infty}\frac{dL_-(t)}{dt}=0,\ L_0=\text{const}. \tag{35′}$$

Тогда согласно (35) и (35′)

$$\lim_{t\to+\infty}L_1(t)=L_0,$$

откуда следует, что

$$|L_0|\neq 0,$$

поскольку определитель $|L_1(t)|$ ограничен по модулю снизу.

Подставляя в (33) вместо $L_1(t)$ сумму $L_-(t)+L_0$, получим:

$$\frac{dL_-(t)}{dt}-A_1L_-(t)+B_1L_-(t)=A_1L_0-B_1L_0,$$

откуда в силу (35′)

$$A_1L_0-L_0B_1=0$$

и, следовательно,

$$B_1=L_0^{-1}A_1L_0. \tag{36}$$

Обратно, если имеет место (36), то преобразование Ляпунова

$$U=L_0V$$

переводит систему (31) в систему (32). Теорема доказана.

Из доказанной теоремы вытекает, что *всякая приводимая система* (19) *при помощи преобразования Ляпунова* $X=LY$ *может быть приведена к виду*

$$\frac{dY}{dt}=JY,$$

если матрицы $A_1-\Lambda_0E$ и B_1 имеют общее характеристическое число, то $\Lambda_0=\gamma-\delta$ где γ и δ — характеристические числа соответственно матриц A_1 и B_1. Подробное исследование оператора $\widehat{F}$ можно найти в работах Ф. Голубчикова [85а, б].

где J — жорданова матрица с вещественными характеристическими числами. Эта каноническая форма системы заданием матрицы $P(t)$ определяется однозначно с точностью до порядка диагональных клеток в J.

§ 5. Матрицант

Рассмотрим систему дифференциальных уравнений

$$\frac{dX}{dt} = P(t)\,X, \tag{37}$$

где $P(t) = \| p_{ik}(t) \|_1^n$ — непрерывная матричная функция в некотором интервале (a, b) изменения аргумента t [1]).

Воспользуемся методом последовательных приближений для определения нормированного решения системы (37), т. е. решения, обращающегося в единичную матрицу при $t = t_0$ [t_0 — фиксированное число из интервала (a, b)]. Последовательные приближения X_k $(k = 0, 1, 2, \ldots)$ будем находить из рекуррентных соотношений

$$\frac{dX_k}{dt} = P(t)\,X_{k-1} \qquad (k = 1, 2, \ldots),$$

выбирая в качестве приближения X_0 единичную матрицу E.

Полагая $X_k(t_0) = E$ $(k = 0, 1, 2, \ldots)$, мы сможем X_k представить в виде

$$X_k = E + \int_{t_0}^{t} P(t)\,X_{k-1}\,dt \qquad (k = 1, 2, \ldots).$$

Таким образом,

$$X_0 = E,\ X_1 = E + \int_{t_0}^{t} P(t)\,dt,\ X_2 = E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt,\ \ldots,$$

т. е. X_k $(k = 0, 1, 2, \ldots)$ есть сумма первых $k+1$ членов матричного ряда

$$E + \int_{t_0}^{t} P(t)\,dt + \int_{t_0}^{t} P(t)\,dt \int_{t_0}^{t} P(t)\,dt + \ldots \tag{38}$$

Для того чтобы доказать, что этот ряд абсолютно и равномерно сходится в любой замкнутой части интервала (a, b) и определяет искомое решение уравнения (37), мы построим мажорантный ряд.

Определим неотрицательные функции $g(t)$ и $h(t)$ в интервале (a, b) равенствами [2])

$$g(t) = \max\left[\,|p_{11}(t)|,\ |p_{12}(t)|,\ \ldots,\ |p_{nn}(t)|\,\right],\quad h(t) = \left|\int_{t_0}^{t} g(t)\,dt\right|.$$

---

[1]) (a, b) — произвольный интервал (конечный или бесконечный). Все элементы $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ матрицы $P(t)$ — комплексные функции вещественного аргумента t, непрерывные в интервале (a, b). Все последующее сохраняет силу, если вместо непрерывности потребовать лишь ограниченность и интегрируемость по Риману [в любом конечном подинтервале интервала (a, b)] всех функций $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$.

[2]) По определению значение функции $g(t)$ при каком-либо из значений t равно наибольшему из n^2 модулей значений $p_{ik}(t)$ $(i, k = 1, 2, \ldots, n)$ при том же значении t.

Легко проверяется, что функции $g(t)$, а следовательно, и $h(t)$ непрерывны в интервале (a, b)[1]).

Каждый из n^2 скалярных рядов, на которые распадается матричный ряд (38), мажорируется рядом

$$1+h(t)+\frac{nh^2(t)}{2!}+\frac{n^2h^3(t)}{3!}+\ldots \tag{39}$$

Действительно,

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right|=\left|\int_{t_0}^{t} p_{ik}(t)\,dt\right|\leqslant\left|\int_{t_0}^{t} g(t)\,dt\right|=h(t),$$

$$\left|\left\{\int_{t_0}^{t} P(t)\,dt\int_{t_0}^{t} P(t)\,dt\right\}_{i,k}\right|=\left|\sum_{j=1}^{n}\int_{t_0}^{t} p_{ij}\,dt\int_{t_0}^{t} p_{jk}(t)\,dt\right|\leqslant$$

$$\leqslant n\left|\int_{t_0}^{t} g(t)\,dt\int_{t_0}^{t} g(t)\,dt\right|=\frac{nh^2(t)}{2},$$

и т. д.

Ряд (39) сходится в интервале (a, b), причем сходится равномерно в любой замкнутой части этого интервала. Отсюда вытекает, что и матричный ряд (38) сходится в (a, b) и притом абсолютно и равномерно в любом замкнутом интервале, входящем в (a, b).

Почленным дифференцированием проверяем, что сумма ряда (38) представляет собой решение уравнения (37); это решение обращается в E при $t=t_0$. Почленное дифференцирование ряда (38) допустимо, поскольку ряд, получающийся после дифференцирования, отличается множителем P от ряда (38) и, следовательно, как и ряд (38), является равномерно сходящимся в любой замкнутой части интервала (a, b).

Таким образом, нами доказана теорема о существовании нормированного решения уравнения (37). Это решение будем обозначать через $\Omega_{t_0}^{t}(P)$ или просто $\Omega_{t_0}^{t}$. Любое другое решение, как было показано § 1, имеет вид

$$X=\Omega_{t_0}^{t}C,$$

где C — произвольная постоянная матрица. Из этой формулы следует, что любое решение, и в частности нормированное, однозначно определяется своим значением при $t=t_0$.

Нормированное решение $\Omega_{t_0}^{t}$ уравнения (37) часто называют *матрицантом*.

Мы показали, что матрицант представим в виде ряда[2])

$$\Omega_{t_0}^{t}=E+\int_{t_0}^{t} P(t)\,dt+\int_{t_0}^{t} P(t)\,dt\int_{t_0}^{t} P(t)\,dt+\ldots, \tag{40}$$

который сходится абсолютно и равномерно в любом замкнутом интервале, в котором функция $P(t)$ непрерывна.

---

[1]) Непрерывность функции $g(t)$ в любой точке t_1 интервала (a, b) следует из того, что разность $g(t)-g(t_1)$ при t, достаточно близком к t_1, всегда совпадает с одной из n^2 разностей $|p_{ik}(t)|-|p_{ik}(t_1)|$ $(i, k=1, 2, \ldots, n)$.

[2]) Представление матрицанта в виде такого ряда было впервые получено Пеано [226].

Отметим некоторые формулы для матрицанта.

1° $$\Omega_{t_0}^t = \Omega_{t_1}^t \Omega_{t_0}^{t_1} \quad (t_0,\ t_1,\ t \subset (a,\ b)).$$

Действительно, поскольку $\Omega_{t_0}^t$ и $\Omega_{t_1}^t$ — два решения уравнения (37), то

$$\Omega_{t_0}^t = \Omega_{t_1}^t C \quad (C \text{ — постоянная матрица}).$$

Полагая здесь $t = t_1$, получим $C = \Omega_{t_0}^{t_1}$.

2° $\Omega_{t_0}^t (P+Q) = \Omega_{t_0}^t (P)\, \Omega_{t_0}^t (S)$, где $S = [\Omega_{t_0}^t (P)]^{-1} Q \Omega_{t_0}^t (P)$.

Для вывода этой формулы положим:

$$X = \Omega_{t_0}^t (P), \quad Y = \Omega_{t_0}^t (P+Q)$$

и

$$Y = XZ. \tag{41}$$

Дифференцируя почленно (41), найдем:

$$(P+Q)\, XZ = PXZ + X \frac{dZ}{dt}.$$

Отсюда

$$\frac{dZ}{dt} = X^{-1} Q X Z$$

и, следовательно, поскольку из (41) следует, что $Z(t_0) = E$,

$$Z = \Omega_{t_0}^t (X^{-1} Q X).$$

Подставляя в (41) вместо X, Y, Z соответствующие матрицанты, получаем формулу 2°.

3° $$\ln |\Omega_{t_0}^t (P)| = \int_{t_0}^{t} \operatorname{Sp} P\, dt.$$

Эта формула следует из тождества Якоби (4) (стр. 420), если в него вместо $X(t)$ подставить $\Omega_{t_0}^t (P)$.

4° *Если* $A = \| a_{ik} \|_1^n = \text{const}$, *то*

$$\Omega_{t_0}^t (A) = e^{A(t-t_0)}.$$

Введем следующие обозначения. Если $P = \| p_{ik} \|_1^n$, то через $\operatorname{mod} P$ будем обозначать матрицу

$$\operatorname{mod} P = \| \, | p_{ik} | \, \|_1^n.$$

Кроме того, если $A = \| a_{ik} \|_1^n$ и $B = \| b_{ik} \|_1^n$ — две вещественные матрицы и

$$a_{ik} \leqslant b_{ik} \quad (i,\ k = 1,\ 2,\ \ldots,\ n),$$

то мы будем писать:

$$A \leqslant B.$$

Тогда из представления (40) следует:

5° *Если* $\operatorname{mod} P(t) \leqslant Q(t)$, *то*

$$\operatorname{mod} \Omega_{t_0}^t (P) \leqslant \Omega_{t_0}^t (Q) \quad (t > t_0).$$

В дальнейшем матрицу n-го порядка, у которой все элементы равны единице, будем обозначать через I:

$$I = \| 1 \|.$$

Рассмотрим функцию $g(t)$, определенную на стр. 429. Тогда

$$\operatorname{mod} P(t) \leqslant g(t) I.$$

Отсюда в силу 5°

$$\operatorname{mod} \Omega_{t_0}^t(P) \leqslant \Omega_{t_0}^t(g(t) I) \quad (t > t_0). \tag{42}$$

Но $\Omega_{t_0}^t(g(t) I)$ есть нормированное решение уравнения

$$\frac{dX}{dt} = g(t) IX.$$

Следовательно, в силу 4°[1])

$$\Omega_{t_0}^t(g(t) I) = e^{h(t) I} \leqslant \left(1 + h(t) + \frac{nh^2(t)}{2!} + \frac{n^2h^3(t)}{3!} + \ldots\right) I,$$

где

$$h(t) = \int_{t_0}^t g(t)\, dt.$$

Поэтому из (42) следует:

6° $$\operatorname{mod} \Omega_{t_0}^t(P) \leqslant \left(\frac{1}{n} e^{nh(t)} + \frac{n-1}{n}\right) I \leqslant e^{nh(t)} I \quad (t > t_0),$$

где

$$h(t) = \int_{t_0}^t g(t)\, dt, \quad g(t) = \max_{1 \leqslant i,\, k \leqslant n} \{|p_{ik}(t)|\}.$$

Покажем теперь, как при помощи матрицанта выражается общее решение системы линейных дифференциальных уравнений с правыми частями:

$$\frac{dx_i}{dt} = \sum_{k=1}^n p_{ik}(t) x_k + f_i(t) \quad (i = 1, 2, \ldots, n); \tag{43}$$

$p_{ik}(t)$, $f_i(t)$ $(i, k = 1, 2, \ldots, n)$ — непрерывные функции в интервале изменения аргумента t.

Вводя столбцевые матрицы («векторы») $x = (x_1, x_2, \ldots, x_n)$ и $f = (f_1, f_2, \ldots, f_n)$ и квадратную матрицу $P = \| p_{ik} \|_1^n$, запишем эту систему так:

$$\frac{dx}{dt} = P(t) x + f(t). \tag{43'}$$

Будем искать решение этого уравнения в виде

$$x = \Omega_{t_0}^t(P) z, \tag{44}$$

где z — неизвестный столбец, зависящий от t. Подставим это выражение для x в (43'), получим:

$$p\Omega_{t_0}^t(P) z + \Omega_{t_0}^t(P) \frac{dz}{dt} = P\Omega_{t_0}^t(P) z + f(t),$$

откуда

$$\frac{dz}{dt} = [\Omega_{t_0}^t(P)]^{-1} f(t).$$

---

[1]) Используя замену независимой переменной t переменной $h = \int_{t_0}^t g(t)\, dt$.

Интегрируя, находим:

$$z = \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + c,$$

где c — произвольный постоянный вектор. Подставим это выражение в (44), получим:

$$x = \Omega_{t_0}^{t}(P) \int_{t_0}^{t} [\Omega_{t_0}^{\tau}(P)]^{-1} f(\tau)\, d\tau + \Omega_{t_0}^{t}(P)\, c. \tag{45}$$

Давая t значение t_0, найдем: $x(t_0) = c$. Поэтому формула (45) принимает вид

$$x = \Omega_{t_0}^{t}(P)\, x(t_0) + \int_{t_0}^{t} K(t, \tau) f(\tau)\, d\tau, \tag{45'}$$

где

$$K(t, \tau) = \Omega_{t_0}^{t}(P)\, [\Omega_{t_0}^{\tau}(P)]^{-1}$$

— так называемая матрица Коши.

§ 6. Мультипликативный интеграл. Инфинитезимальное исчисление Вольтерра

Рассмотрим матрицант $\Omega_{t_0}^{t}(P)$. Разобьем основной интервал (t_0, t) на n частей, введя промежуточные точки $t_1, t_2, \ldots, t_{n-1}$, и положим $\Delta t_k = t_k - t_{k-1}$ $(k = 1, 2, \ldots, n;\ t_n = t)$. Тогда на основании свойства 1° матрицанта (см. предыдущий параграф)

$$\Omega_{t_0}^{t} = \Omega_{t_{n-1}}^{t} \ldots \Omega_{t_1}^{t_2} \Omega_{t_0}^{t_1}. \tag{46}$$

Выберем в интервале (t_{k-1}, t_k) промежуточную точку τ_k $(k = 1, 2, \ldots, n)$. Тогда, считая Δt_k малыми величинами первого порядка, при вычислении $\Omega_{t_{k-1}}^{t_k}$ с точностью до малых второго порядка можно принять $P(t) \approx \text{const} = P(\tau_k)$. Тогда

$$\Omega_{t_{k-1}}^{t_k} = e^{P(\tau_k)\Delta t_k} + (**) = E + P(\tau_k)\Delta t_k + (**); \tag{47}$$

здесь символом $(**)$ мы обозначаем сумму членов, начиная со второго порядка малости.

Из (46) и (47) находим:

$$\Omega_{t_0}^{t} = e^{P(\tau_n)\Delta t_n} \ldots e^{P(\tau_2)\Delta t_2} e^{P(\tau_1)\Delta t_1} + (*) \tag{48}$$

и

$$\Omega_{t_0}^{t} = [E + P(\tau_n)\Delta t_n] \ldots [E + P(\tau_2)\Delta t_2]\, [E + P(\tau_1)\Delta t_1] + (*). \tag{49}$$

Переходя к пределу при неограниченном увеличении числа интервалов разбиения и стремлении к нулю длин этих интервалов (при предельном переходе малые члены $(*)$ исчезают[1]), получаем точные предельные

[1]) Эти рассуждения могут быть уточнены путем оценки членов, обозначенных нами через $(*)$.

формулы

$$\Omega_{t_0}^{t}(P) = \lim_{\Delta t_k \to 0} [e^{P(\tau_n)\Delta t_n} \ldots e^{P(\tau_2)\Delta t_2} e^{P(\tau_1)\Delta t_1}] \tag{48'}$$

и

$$\Omega_{t_0}^{t}(P) = \lim_{\Delta t_k \to 0} [E + P(\tau_n)\Delta t_n] \ldots [E + P(\tau_2)\Delta t_2][E + P(\tau_1)\Delta t_1]. \tag{49'}$$

Выражение, стоящее под знаком предела в правой части последнего равенства, представляет собой *интегральное произведение*[1]). Предел его мы назовем *мультипликативным интегралом* и обозначим символом

$$\widehat{\int_{t_0}^{t}} [E + P(t)\,dt] = \lim_{\Delta t_k \to 0} [E + P(\tau_n)\Delta t_n] \ldots [E + P(\tau_1)\Delta t_1]. \tag{50}$$

Формула (49′) дает представление матрицанта в виде мультипликативного интеграла

$$\Omega_{t_0}^{t}(P) = \widehat{\int_{t_0}^{t}} (E + P\,dt), \tag{51}$$

а равенства (48) и (49) могут быть использованы для приближенного вычисления матрицанта.

Мультипликативный интеграл ввел впервые Вольтерра в 1887 г. На базе этого понятия Вольтерра построил своеобразное инфинитезимальное исчисление для матричных функций (см. [49])[2]).

Вся специфика мультипликативного интеграла связана с неперестановочностью между собой различных значений подинтегральной матричной функции $P(t)$. В том же весьма частном случае, когда все эти значения перестановочны между собой

$$P(t')P(t'') = P(t'')P(t') \qquad (t', t'' \in (t_0, t)),$$

мультипликативный интеграл, как это видно из (48′) и (51), сводится к матрице

$$e^{\int_{t_0}^{t} P(t)\,dt}.$$

Введем теперь *мультипликативную производную*

$$D_t X = \frac{dX}{dt} X^{-1}. \tag{52}$$

Операции D_t и $\widehat{\int_{t_0}^{t}}$ взаимно обратны:

Если

$$D_t X = P,$$

---

[1]) Аналог интегральной суммы для обычного интеграла.

[2]) Мультипликативный интеграл (по-немецки «Produkt-Integral») был использован Шлезингером при исследовании систем линейных дифференциальных уравнений с аналитическими коэффициентами [61a, b], см. также [235].

Мультипликативный интеграл (50) существует не только для функции $P(t)$, непрерывной в интервале интегрирования, но и при значительно более общих предположениях (см. [154]).

то

$$X = \int_{t_0}^{t} (E + P\,dt) \cdot C \qquad (C = X(t_0))\,^{1)}$$

и наоборот. Последняя формула может быть еще записана так [2]):

$$\int_{t_0}^{t} (E + P\,dt) = X(t)\,X(t_0)^{-1}. \tag{53}$$

Предлагаем читателю проверить справедливость следующих дифференциальных и интегральных формул [3]):

Дифференциальные формулы

I. $D_t(XY) = D_t(X) + XD_t(Y)\,X^{-1}$,
$D_t(XC) = D_t(X)$, $\left(C \text{ — постоянная матрица}\right)$.
$D_t(CY) = CD_t(Y)\,C^{-1}$

II. $D_t(X') = X'\,(D_tX)'X'^{-1}$ [4]).

III. $D_t(X^{-1}) = -X^{-1}D_t(X)\,X = -(D_t(X'))'$,
$D_t(X'^{-1}) = -(D_t(X))'$.

Интегральные формулы

IV. $$\int_{t_0}^{t} (E + P\,dt) = \int_{t_1}^{t} (E + P\,dt) \int_{t_0}^{t_1} (E + P\,dt).$$

V. $$\int_{t_0}^{t} (E + P\,dt) = \left[\int_{t}^{t_0} (E + P\,dt)\right]^{-1}.$$

VI. $$\int_{t_0}^{t} (E + CPC^{-1}\,dt) = C \int_{t_0}^{t} (E + P\,dt)\,C^{-1} \quad \left(C \text{ — постоянная матрица}\right).$$

VII. $$\int_{t_0}^{t} [E + (Q + D_tX)\,dt] = X(t) \int_{t_0}^{t} (E + X^{-1}QX\,dt)\,X(t_0)^{-1}\,^{5)}.$$

---

1) Здесь произвольная постоянная матрица C является аналогом аддитивной произвольной постоянной в обычном неопределенном интеграле.

2) Аналог формулы $\int_{t_0}^{t} P\,dt = X(t) - X(t_0)$ в случае, когда $\frac{dX}{dt} = P$.

3) Эти формулы могут быть выведены непосредственно из определения мультипликативных производной и интеграла (см. [67]). Однако интегральные формулы получаются быстрее и проще, если рассматривать мультипликативный интеграл как матрицант и воспользоваться свойствами матрицанта, изложенными в предыдущем параграфе (см. [61a]).

4) Значок $'$ обозначает переход к транспонированной матрице.

5) Формула VII может быть рассматриваема в известном смысле как аналог формулы интегрирования по частям для обычных (немультипликативных) интегралов. Формула VII следует из формулы 2° § 5.

Выведем еще важную формулу, дающую оценку модуля[1]) разности между двумя мультипликативными интегралами:

$$\text{VIII. } \operatorname{mod}\left[\widehat{\int}_{t_0}^{t}(E+P\,dt)-\widehat{\int}_{t_0}^{t}(E+Q\,dt)\right]\leqslant\frac{1}{n}\,e^{nq(t-t_0)}(e^{nd(t-t_0)}-1)I(t>t_0),$$

если

$$\operatorname{mod} Q\leqslant qI,\qquad \operatorname{mod}(P-Q)\leqslant d\cdot I,\qquad I=\|1\|$$

(q, d — неотрицательные числа, n — порядок матриц P и Q).

Обозначим через D разность $P-Q$. Тогда

$$P=Q+D,\qquad \operatorname{mod} D\leqslant d\cdot I.$$

Рассматривая мультипликативный интеграл как матрицант и пользуясь разложением (40) матрицанта в ряд, найдем:

$$\widehat{\int}_{t_0}^{t}[E+(Q+D)\,dt]-\widehat{\int}_{t_0}^{t}(E+Q\,dt)=$$
$$=\int_{t_0}^{t}D\,dt+\int_{t_0}^{t}D\,dt\int_{t_0}^{t}Q\,dt+\int_{t_0}^{t}Q\,dt\int_{t_0}^{t}D\,dt+\int_{t_0}^{t}D\,dt\int_{t_0}^{t}D\,dt+\ldots$$

Из этого разложения видно, что

$$\operatorname{mod}\left\{\widehat{\int}_{t_0}^{t}[E+(Q+D)\,dt]-\widehat{\int}_{t_0}^{t}(E+Q\,dt)\right\}\leqslant$$
$$\leqslant\widehat{\int}_{t_0}^{t}E+(\operatorname{mod}Q+\operatorname{mod}D)\,dt-\widehat{\int}_{t_0}^{t}(E+\operatorname{mod}Q)\,dt\leqslant$$
$$\leqslant\widehat{\int}_{t_0}^{t}[E+(q+d)\,I\,dt]-\widehat{\int}_{t_0}^{t}[E+qI]\,dt=e^{(q+d)I(t-t_0)}-e^{qI(t-t_0)}=$$
$$=e^{qI(t-t_0)}\,(e^{d\cdot I(t-t_0)}-E)\leqslant\frac{1}{n}\,e^{nq(t-t_0)}[e^{nd(t-t_0)}-1]\,I.$$

Пусть теперь матрицы P и Q зависят от некоторого параметра α:

$$P=P(t,\ \alpha),\qquad Q=Q(t,\ \alpha),$$

и пусть

$$\lim_{\alpha\to\alpha_0}P(t,\ \alpha)=\lim_{\alpha\to\alpha_0}Q(t,\ \alpha)=P_0(t),$$

причем стремление к пределу равномерно относительно t в рассматриваемом интервале $(t_0,\ t)$. Допустим, что, кроме того, при $\alpha\to\alpha_0$ матрица $Q(t,\ \alpha)$ по модулю ограничена матрицей qI, где q — положительная постоянная. Тогда, полагая

$$d(\alpha)=\max_{\substack{1\leqslant i,\ k\leqslant n\\ t_0\leqslant\tau\leqslant t}}|p_{ik}(\tau,\ \alpha)-q_{ik}(\tau,\ \alpha)|,$$

будем иметь:

$$\lim_{\alpha\to\alpha_0}d(\alpha)=0.$$

---

[1]) Относительно определения модуля матрицы, а также соотношения $\leqslant$ между матрицами см. стр. 431.

Поэтому из формулы VIII следует:

$$\lim_{\alpha\to\alpha_0}\left[\overset{\frown}{\int_{t_0}^{t}}(E+P\,dt)-\overset{\frown}{\int_{t_0}^{t}}(E+Q\,dt)\right]=0.$$

В частности, если Q не зависит от α $[Q(t,\alpha)=P_0(t)]$, мы получаем:

$$\lim_{\alpha\to\alpha_0}\overset{\frown}{\int_{t_0}^{t}}[E+P(t,\alpha)\,dt]=\overset{\frown}{\int_{t_0}^{t}}[E+P_0(t)\,dt],$$

где

$$P_0(t)=\lim_{\alpha\to\alpha_0}P(t,\alpha).$$

§ 7. Дифференциальные системы в комплексной области. Общие свойства

Рассмотрим систему дифференциальных уравнений

$$\frac{dx_i}{dz}=\sum_{k=1}^{n}p_{ik}(z)\,x_k. \tag{54}$$

Здесь данные функции $p_{ik}(z)$ и искомые функции $x_i(z)$ $(i,\ k=1,\ 2,\ \ldots,\ n)$ предполагаются однозначными аналитическими функциями комплексного аргумента z, регулярными в некоторой области G комплексной z-плоскости.

Вводя квадратную матрицу $P(z)=\|p_{ik}(z)\|_1^n$ и столбцевую матрицу $x=(x_1,\ x_2,\ \ldots,\ x_n)$, мы, как и в случае вещественного аргумента (§ 1), можем записать систему (54) в виде

$$\frac{dx}{dz}=P(z)\,x. \tag{54'}$$

Обозначая через X интегральную матрицу, т. е. матрицу, столбцами которой являются n линейно независимых решений системы (54), мы вместо (54′) можем написать:

$$\frac{dX}{dz}=P(z)\,X. \tag{55}$$

Формула Якоби имеет место и при комплексном аргументе z:

$$|X|=ce^{\int_{z_0}^{z}\operatorname{Sp}P\,dz} \tag{56}$$

При этом предполагается, что z_0 и все точки пути, вдоль которого берется $\int_{z_0}^{z}$, являются регулярными точками для однозначной аналитической функции $\operatorname{Sp}P(z)=p_{11}(z)+p_{22}(z)+\ldots+p_{nn}(z)$ [1]).

Специфичность рассматриваемого случая комплексного аргумента заключается в том, что при однозначной функции $P(z)$ интегральная матрица $X(z)$ может быть многозначной функцией от z.

---

[1]) Здесь и в дальнейшем в качестве путей интегрирования берутся кусочно гладкие кривые.

В качестве примера рассмотрим систему Коши

$$\frac{dX}{dz}=\frac{U}{z-a}X \quad (U\text{ — постоянная матрица}). \tag{57}$$

Одним из решений этой системы, как и в случае вещественного аргумента, является (см. стр. 421) интегральная матрица

$$X=e^{U\ln(z-a)}=(z-a)^U. \tag{58}$$

В качестве области G возьмем всю z-плоскость за исключением точки $z=a$. Все точки этой области являются регулярными точками матрицы коэффициентов

$$P(z)=\frac{U}{z-a}.$$

Если $U\neq 0$, то точка $z=a$ является особой точкой (полюсом первого порядка) для матричной функции $P(z)=\frac{U}{z-a}$.

Элемент интегральной матрицы (58) при однократном обходе в положительном направлении точки $z=a$ возвращается с новым значением, которое получается из старого умножением справа на постоянную матрицу

$$V=e^{2\pi iU}.$$

Для общей системы (55) теми же рассуждениями, что и в случае вещественного аргумента, убеждаемся в том, что два однозначных решения X и $\widetilde{X}$ в некоторой части области G всегда связаны формулой

$$X=\widetilde{X}C,$$

где C — некоторая постоянная матрица. Эта формула сохранится при любом аналитическом продолжении функций $X(z)$ и $\widetilde{X}(z)$ в области G.

Теорема о существовании и единственности (при заданных начальных значениях) решения системы (54) может быть доказана аналогично вещественному случаю.

Рассмотрим односвязную и притом звездообразную относительно точки z_0[1]) область G_1, составляющую часть области G, и пусть матричная функция $P(z)$ регулярна[2]) в области G_1. Составим ряд

$$E+\int_{z_0}^{z}P\,dz+\int_{z_0}^{z}P\,dz\int_{z_0}^{z}P\,dz+\ldots \tag{59}$$

Из односвязности области G_1 следует, что каждый встречающийся в ряду (59) интеграл не зависит от пути интегрирования и представляет собой регулярную функцию в области G_1. Поскольку область G_1 звездообразна относительно z_0, то при оценке модулей этих интегралов мы можем считать, что все интегралы берутся вдоль прямолинейного отрезка, соединяющего точки z_0 и z.

Абсолютная и равномерная в любой замкнутой части области G_1, содержащей точку z_0, сходимость ряда (59) вытекает из сходимости

---

[1]) Область называется *звездообразной относительно точки* z_0, если любой отрезок, соединяющий произвольную точку z области с точкой z_0, целиком лежит в данной области.

[2]) То есть все элементы $p_{ik}(z)$ $(i, k=1, 2, \ldots n)$ матрицы $P(z)$ являются регулярными функциями в области G_1.

мажорантного ряда

$$1+lM+\frac{n}{2!}l^2M^2+\frac{n^2}{3!}l^3M^3+\ldots$$

Здесь M — верхняя граница для модуля матрицы $P(z)$, а l — верхняя граница расстояний точки z от точки z_0, причем обе границы относятся к рассматриваемой замкнутой части области G_1.

Путем почленного дифференцирования проверяется, что сумма ряда (59) представляет собой решение уравнения (55). Это решение нормировано, поскольку оно при $z=z_0$ обращается в единичную матрицу E. Однозначное нормированное решение системы (55), как и в вещественном случае, будем называть матрицантом и будем обозначать через $\Omega_{z_0}^z(P)$. Таким образом, мы получили представление матрицанта в области G_1 в виде ряда[1])

$$\Omega_{z_0}^z(P)=E+\int_{z_0}^{z}P\,dz+\int_{z_0}^{z}P\,dz\int_{z_0}^{z}P\,dz+\ldots \tag{60}$$

Свойства 1°—4° *матрицанта, установленные в* § 5, *автоматически переносятся и на случай комплексного аргумента.*

Произвольное решение уравнения (55), регулярное в области G и обращающееся при $z=z_0$ в матрицу X_0, представится в виде

$$X=\Omega_{z_0}^z(P)\cdot C \qquad (C=X_0). \tag{61}$$

Формула (61) охватывает все однозначные решения, регулярные в окрестности точки z_0 [z_0 — регулярная точка для матрицы коэффициентов $P(z)$]. Эти решения, будучи аналитически продолжены в область G, дадут все решения уравнения (55), т. е. уравнение (55) не может иметь решений, для которых z_0 была бы особой точкой.

Для аналитического продолжения матрицанта в область G удобно пользоваться мультипликативным интегралом.

§ 8. Мультипликативный интеграл в комплексной области

Мультипликативный интеграл вдоль некоторой кривой в комплексной плоскости определяется следующим образом.

Пусть даны некоторый путь L и матричная функция $P(z)$, непрерывная на кривой L. Разобьем путь L на n частей $(z_0, z_1)(z_1, z_2)\ldots(z_{n-1}, z)$; здесь z_0 — начало, $z_n=z$ — конец пути, а $z_1, z_2, \ldots, z_{n-1}$ — промежуточные точки разбиения. На отрезке пути (z_{k-1}, z_k) выберем произвольную точку ζ_k и введем обозначения $\Delta z_k=z_k-z_{k-1}$; $k=1, 2, \ldots, n$. Тогда по определению

$$\overset{\frown}{\int_L}[E+P(z)\,dz]=\lim_{\Delta z_k\to 0}[E+P(\zeta_n)\Delta z_n]\ldots[E+P(\zeta_1)\Delta z_1].$$

Сопоставляя это определение с определением на стр. 434, мы видим, что новое определение совпадает с прежним в том частном случае, когда путь L является отрезком вещественной оси. Однако и в общем случае,

---

[1]) Приведенное доказательство существования нормированного решения и представления его в области G_1 рядом (60) сохраняет свою силу, если вместо звездообразности области сделать более общее допущение: для любой замкнутой части области G_1 существует такое положительное число l, что любую точку z этой замкнутой части можно соединить с z_0 путем, длина которого $\leqslant l$.

когда путь L произвольно расположен в комплексной плоскости, новое определение может быть сведено к старому при помощи замены переменной интегрирования.

Если

$$z = z(t)$$

есть параметрическое уравнение пути, причем $z(t)$ — непрерывная функция в интервале (t_0, t), имеющая в этом интервале кусочно непрерывную производную $\frac{dz}{dt}$, то, как легко видеть,

$$\overset{\frown}{\int_L} [E + P(z)\,dz] = \overset{\frown}{\int_{t_0}^{t}} \left\{E + P[z(t)]\frac{dz}{dt}\,dt\right\}.$$

Эта формула показывает, что мультипликативный интеграл вдоль произвольного пути существует, если подинтегральная матрица $P(z)$ непрерывна вдоль этого пути[1]).

Мультипликативная производная определяется прежней формулой

$$D_z X = \frac{dX}{dz} X^{-1}.$$

При этом предполагается, что $X(z)$ — аналитическая функция.

Все дифференциальные формулы (I — III) предыдущего параграфа переносятся без изменения на случай комплексного аргумента. Что же касается интегральных формул IV — VI, то несколько видоизменяется их внешняя запись:

IV'. $$\overset{\frown}{\int_{(L'+L'')}} (E + P\,dz) = \overset{\frown}{\int_{L''}} (E + P\,dz) \overset{\frown}{\int_{L'}} (E + P\,dz).$$

V'. $$\overset{\frown}{\int_{-L}} (E + P\,dz) = \left[\overset{\frown}{\int_{L}} (E + P\,dz)\right]^{-1}.$$

VI'. $$\overset{\frown}{\int_{L}} (E + CPC^{-1}\,dz) = C \overset{\frown}{\int_{L}} (E + P\,dz)\,C^{-1} \left(\begin{matrix} C \text{ — постоянная} \\ \text{матрица} \end{matrix}\right).$$

В формуле IV' мы через $L' + L''$ обозначили составной путь, получающийся при прохождении сначала пути L', а затем пути L''. В формуле V' $-L$ обозначает путь, отличающийся от пути L только направлением.

Формула VII принимает теперь вид

VII'. $$\overset{\frown}{\int_{L}} [E + (Q + D_z X)\,dz] = X(z) \overset{\frown}{\int_{L}} (E + X^{-1}QX\,dz)\,X(z_0)^{-1}.$$

Здесь $X(z_0)$ и $X(z)$ в правой части обозначают соответственно значения $X(z)$ в начале и в конце пути L.

---

[1]) См. сноску [2]) на стр. 434. Даже в случае, когда $P(z)$ — непрерывная функция вдоль L, функция $P[z(t)]\frac{dz}{dt}$ может быть кусочно непрерывной. В этом случае мы можем интервал (t_0, t) разбить на частичные интервалы, в каждом из которых производная $\frac{dz}{dt}$ непрерывна, и под интегралом от t_0 до t понимать сумму интегралов вдоль этих частичных интервалов.

Формула VIII заменится теперь формулой

$$\text{VIII}'. \quad \operatorname{mod}\left[\int_L (E+P\,dz)-\int_L (E+Q\,dz)\right] \leqslant \frac{1}{n} e^{nql}(e^{nd\cdot l}-1)\,I,$$

где $\operatorname{mod} Q \leqslant qI$, $\operatorname{mod}(P-Q) \leqslant d\cdot I$, $I=\|1\|$, а l — длина пути L.

Формула VIII′ получается сразу из формулы VIII, если в последней сделать преобразование переменной, выбрав в качестве новой переменной интегрирования длину дуги s вдоль пути L $\left(\text{при этом } \left|\frac{dz}{ds}\right|=1\right)$.

Как и в случае вещественного аргумента, существует тесная связь мультипликативного интеграла с матрицантом.

Пусть дана однозначная аналитическая матричная функция $P(z)$, регулярная в области G, и пусть G_0 — односвязная область, содержащая точку z_0 и составляющая часть области G. Тогда матрицант $\Omega_{z_0}^{z}(P)$ будет регулярной функцией от z в области G_0.

Соединим точки z_0 и z произвольным путем L, целиком лежащим в G_0, и выберем на L промежуточные точки $z_1, z_2, \ldots, z_{n-1}$. Тогда, пользуясь равенством

$$\Omega_{z_0}^{z}=\Omega_{z_{n-1}}^{z}\ldots\Omega_{z_1}^{z_2}\Omega_{z_0}^{z_1},$$

совершенно так же, как в § 6 (стр. 433), предельным переходом получим:

$$\Omega_{z_0}^{z}(P)=\widehat{\int_L}(E+P\,dz)=\widehat{\int_{z_0}^{z}}(E+P\,dz). \tag{62}$$

Из этой формулы видно, что мультипликативный интеграл не зависит от формы пути, а зависит только от начала и конца пути, если весь путь интегрирования лежит в односвязной области G_0, в которой подинтегральная функция $P(z)$ регулярна. В частности, для замкнутого контура L, лежащего в односвязной области G_0, имеем:

$$\oint(E+P\,dz)=E. \tag{63}$$

Эта формула представляет собой аналог известной теоремы Коши, согласно которой обычный (немультипликативный) интеграл вдоль замкнутого контура равен нулю, если этот контур лежит в односвязной области, в которой подинтегральная функция регулярна.

Представление матрицанта в виде мультипликативного интеграла (62) может быть использовано для аналитического продолжения матрицанта вдоль произвольного пути L в области G. В этом случае формула

$$X=\widehat{\int_{z_0}^{z}}(E+P\,dz)\,X_0 \tag{64}$$

задает все ветви многозначной интегральной матрицы X дифференциального уравнения $\frac{dX}{dz}=PX$, обращающейся в X_0 на одной из ветвей при $z=z_0$. Различные ветви получаются за счет различных путей, соединяющих точки z_0 и z.

Согласно формуле Якоби (56)

$$|X|=|X_0|\,e^{\int_{z_0}^{z}\operatorname{Sp}P\,dz}$$

и, в частности, при $X_0 = E$

$$\left| \widehat{\int_{z_0}^{z}} (E + P\,dz) \right| = e^{\int_{z_0}^{z} \operatorname{Sp} P\,dz}. \tag{65}$$

Из этой формулы следует, что мультипликативный интеграл всегда представляет собой неособенную матрицу, если только путь интегрирования целиком лежит в области, в которой функция $P(z)$ регулярна.

Если L — произвольный замкнутый путь в G и G — неодносвязная область, то равенство (63) может и не иметь места. Более того, в этом случае значение интеграла

$$\widehat{\oint} (E + P\,dz)$$

не определяется заданием подинтегральной функции и замкнутого пути интегрирования L, а зависит еще и от выбора начальной точки интегрирования z_0 на кривой L. Действительно, выберем на замкнутой кривой L две точки z_0 и z_1 и обозначим участки пути от z_0 до z_1 и от z_1 до z_0 (в направлении интегрирования) соответственно через L_1 и L_2. Тогда согласно формуле IV′ [1])

$$\widehat{\oint_{z_0}} = \widehat{\int_{L_2}} \cdot \widehat{\int_{L_1}}, \quad \widehat{\oint_{z_1}} = \widehat{\int_{L_1}} \cdot \widehat{\int_{L_2}}$$

и, следовательно,

$$\widehat{\oint_{z_1}} = \widehat{\int_{L_1}} \cdot \widehat{\oint_{z_0}} \cdot \widehat{\int_{L_1}}^{-1}. \tag{66}$$

Формула (66) показывает, что символ $\widehat{\oint} (E + P\,dz)$ определяет некоторую матрицу с точностью до преобразования подобия, т. е. определяет только элементарные делители некоторой матрицы.

Рассмотрим элемент $X(z)$ решения (64) в окрестности точки z_0. Пусть L — произвольный замкнутый путь в G, начинающийся и кончающийся в точке z_0. После аналитического продолжения вдоль L элемент $X(z)$ перейдет в некоторый элемент $\widetilde{X}(z)$. При этом новый элемент $\widetilde{X}(z)$ будет удовлетворять тому же дифференциальному уравнению (55), поскольку $P(z)$ — однозначная функция в G. Поэтому $\widetilde{X} = XV$, где V — некоторая неособенная постоянная матрица. Из формулы (64) следует, что

$$\widetilde{X}(z_0) = \widehat{\oint_{z_0}} (E + P\,dz)\, X_0.$$

Сопоставляя это равенство с предыдущим, найдем:

$$V = X_0^{-1} \widehat{\oint_{z_0}} (E + P\,dz)\, X_0. \tag{67}$$

В частности, для матрицанта $X = \Omega_z^z$ имеем $X_0 = E$, и тогда

$$V = \widehat{\oint_{z_0}} (E + P\,dz). \tag{68}$$

[1]) Здесь мы для сокращения обозначений опускаем подинтегральное выражение $E + P\,dz$, одно и то же во всех интегралах.

§ 9. Изолированная особая точка

Займемся исследованием поведения решения (интегральной матрицы) в окрестности изолированной особой точки a.

Пусть матричная функция $P(z)$ регулярна для значений z, удовлетворяющих неравенствам

$$0<|z-a|<R.$$

Совокупность этих значений образует двусвязную область G. Матричная функция $P(z)$ в области G разлагается в ряд Лорана

$$P(z)=\sum_{n=-\infty}^{+\infty} P_n (z-a)^n. \tag{69}$$

Элемент $X(z)$ интегральной матрицы после однократного обхода в положительном направлении вокруг a вдоль пути L перейдет в элемент

$$X^+(z)=X(z)V,$$

где V — некоторая постоянная неособенная матрица.

Пусть U — постоянная матрица, связанная с матрицей V равенством

$$V=e^{2\pi i U}. \tag{70}$$

Тогда матричная функция $(z-a)^U$ после обхода вдоль L также переходит в $(z-a)^U V$. Поэтому аналитическая в области G матричная функция

$$F(z)=X(z)(z-a)^{-U} \tag{71}$$

при аналитическом продолжении вдоль L переходит сама в себя (остается неизменной)[1]). Поэтому матричная функция $F(z)$ регулярна в G и разлагается в G в ряд Лорана

$$F(z)=\sum_{n=-\infty}^{+\infty} F_n (z-a)^n. \tag{72}$$

Из (71) следует:

$$X(z)=F(z)(z-a)^U. \tag{73}$$

Таким образом, каждая интегральная матрица $X(z)$ может быть представлена в виде (73), где однозначная функция $F(z)$ и постоянная матрица U зависят от матрицы коэффициентов $P(z)$. Однако алгорифмическое определение матрицы U и коэффициентов F_n ряда (72) по коэффициентам P_n ряда (69) в общем случае представляет собой сложную задачу.

Частный случай этой задачи, когда

$$P(z)=\sum_{n=-1}^{\infty} P_n (z-a)^n,$$

---

[1]) Отсюда уже следует, что функция $F(z)$ при обходе вдоль любого другого замкнутого пути в G возвращается к исходному значению.

будет нами разобран полностью в § 10. В этом случае точка a называется *регулярной особой точкой* системы (55).

Если разложение (69) имеет вид

$$P(z)=\sum_{n=-q}^{\infty} P_n (z-a)^n \qquad (q>1;\ P_{-q}\neq 0),$$

то точка a называется *иррегулярной особой точкой типа полюса*. Наконец, если в ряду (69) имеется бесчисленное множество отличных от нуля матричных коэффициентов P_n при отрицательных степенях $z-a$, то точка a называется *существенной особой точкой* данной дифференциальной системы.

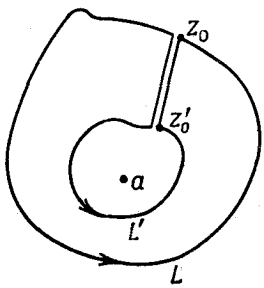

Рис. 8.

Из формулы (73) следует, что интегральная матрица $X(z)$ при любом однократном обходе в положительном направлении (вдоль некоторого замкнутого пути L) умножается справа на одну и ту же матрицу

$$V=e^{2\pi i U}.$$

Если этот обход начинается (и кончается) в точке z_0, то согласно (67)

$$V=X(z_0)^{-1}\oint_{z_0}(E+P\,dz)\,X(z_0). \tag{74}$$

Если вместо интегральной матрицы $X(z)$ мы рассмотрим любую другую интегральную матрицу $\widehat{X}(z)=X(z)C$ (C — постоянная матрица, $|C|\neq 0$), то, как видно из (74), матрица V заменится подобной матрицей

$$\widehat{V}=C^{-1}VC.$$

Таким образом, «интегральные подстановки» V данной системы образуют класс подобных между собой матриц.

Из формулы (74) также следует, что интеграл

$$\oint_{z_0}(E+P\,dz) \tag{75}$$

определяется начальной точкой обхода z_0 и не зависит от формы кривой обхода[1]). Если же мы меняем точку z_0, то получающиеся при этом различные значения интеграла (75) подобны между собой[2]).

В этих свойствах интеграла (75) можно убедиться и непосредственно. Действительно, пусть L и L' — два замкнутых пути в G вокруг точки $z=a$ с начальными точками обхода z_0 и z_0' (см. рис. 8).

Двусвязная область, заключенная между L и L', может быть сделана односвязной, если провести разрез от z_0 до z_0'. Интеграл вдоль разреза мы обозначим через

$$T=\int_{z_0}^{z_0'}(E+P\,dz).$$

---

[1]) Конечно, при условии, что путь интегрирования обходит точку a однократно в положительном направлении.

[2]) Это вытекает из формулы (74), а также из формулы (66).

Поскольку мультипликативный интеграл вдоль замкнутого контура односвязной области равен E, то

$$\widehat{\int_{L'}} T \widehat{\int_{L}}^{-1} T^{-1} = E,$$

откуда

$$\widehat{\int_{L'}} = T \widehat{\int_{L}} T^{-1}.$$

Таким образом, как и V, интеграл $\oint (E + P\,dz)$ определен с точностью до подобия, и равенство (74) мы иногда будем записывать так:

$$V = \widehat{\oint} (E + P\,dz),$$

понимая под этим совпадение элементарных делителей у матриц, стоящих в левой и правой частях равенства.

Рассмотрим для примера систему с регулярной особой точкой

$$\frac{dX}{dz} = P(z)\,X,$$

где

$$P(z) = \frac{P_{-1}}{z-a} + \sum_{n=0}^{+\infty} P_n (z-a)^n.$$

Пусть

$$Q(z) = \frac{P_{-1}}{z-a}.$$

Пользуясь формулой VIII′ предыдущего параграфа, дадим оценку модуля разности

$$D = \widehat{\oint} (E + P\,dz) - \widehat{\oint} (E + Q\,dz), \tag{76}$$

выбрав в качестве пути интегрирования окружность радиуса $r\,(r < R)$ с положительным направлением обхода. Тогда при

$$\operatorname{mod} P_{-1} \leqslant p_{-1} I, \quad \operatorname*{mod}_{|z-a|=r} \sum_{n=0}^{\infty} P_n (z-a)^n \leqslant d(r)\,I, \quad I = \|1\|$$

мы можем в формуле VIII′ положить:

$$q = \frac{p_{-1}}{r}, \quad d = d(r), \quad l = 2\pi r,$$

после чего получим:

$$\operatorname{mod} D \leqslant \frac{1}{n} e^{2\pi p_{-1}} \left(e^{2\pi n r d(r)} - 1\right) I.$$

Отсюда видно, что[1])

$$\lim_{r \to 0} D = 0. \tag{77}$$

---

[1]) При этом мы используем то, что при надлежащем выборе $d(r)$

$$\lim_{r \to 0} d(r) = d_0,$$

где d_0 — наибольший из модулей элементов матрицы P_0.

С другой стороны, система

$$\frac{dY}{dz}=QY$$

является системой Коши, и в этом случае при любом выборе начальной точки обхода z_0 и при любом $r<R$

$$\oint\limits_{z_0}(E+Q\,dz)=e^{2\pi i P_{-1}}.$$

Поэтому из (76) и (77) следует:

$$\lim_{r\to 0}\oint\limits_{z_0}(E+P\,dz)=e^{2\pi i P_{-1}}. \tag{78}$$

Но элементарные делители интеграла $\oint\limits_{z_0}(E+P\,dz)$ не зависят от z_0 и r и совпадают с элементарными делителями интегральной подстановки V.

Отсюда Вольтерра в своем известном мемуаре (см. [149]), а также в книге [49] (стр. 117—120) делает вывод, что матрицы V и $e^{2\pi i P_{-1}}$ подобны, и потому интегральная подстановка V с точностью до подобия определяется матрицей «вычетов» P_{-1}.

Это утверждение Вольтерра ошибочно.

Из (74) и (78) можно лишь сделать вывод, что *характеристические числа интегральной подстановки V совпадают с характеристическими числами матрицы $e^{2\pi i P_{-1}}$*. Однако элементарные делители у этих матриц могут быть различными. Так, например, матрица

$$\begin{Vmatrix}\alpha & r\\ 0 & \alpha\end{Vmatrix}$$

при любом $r\neq 0$ имеет один элементарный делитель $(\lambda-\alpha)^2$, а предел этой матрицы при $r\to 0$, т. е. матрица $\begin{Vmatrix}\alpha & 0\\ 0 & \alpha\end{Vmatrix}$, имеет два элементарных делителя $\lambda-\alpha$, $\lambda-\alpha$.

Таким образом, утверждение Вольтерра не вытекает из формул (74) и (78). Но оно и вообще неверно, как показывает следующий пример.

Пусть

$$P(z)=\begin{Vmatrix}0 & 0\\ 0 & -1\end{Vmatrix}\frac{1}{z}+\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix}.$$

Соответствующая система дифференциальных уравнений имеет вид:

$$\frac{dx_1}{dz}=x_2,\qquad \frac{dx_2}{dz}=-\frac{x_2}{z}.$$

Интегрируя эту систему, находим:

$$x_1=c\ln z+d,\quad x_2=\frac{c}{z}.$$

Интегральная матрица

$$X(z)=\begin{Vmatrix}\ln z & 1\\ z^{-1} & 0\end{Vmatrix}$$

при однократном положительном обходе вокруг особой точки $z=0$ умножается справа на матрицу

$$V=\begin{Vmatrix} 1 & 0 \\ 2\pi i & 1 \end{Vmatrix}.$$

Эта матрица имеет один элементарный делитель $(\lambda-1)^2$. В то же время матрица

$$e^{2\pi i P_{-1}}=e^{2\pi i\begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}}=\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}=E$$

имеет два элементарных делителя $\lambda-1$, $\lambda-1$.

Рассмотрим теперь случай, когда матрица $P(z)$ имеет конечное число отрицательных степеней $z-a$ (a — регулярная или иррегулярная особая точка типа полюса):

$$P(z)=\frac{P_{-q}}{(z-a)^q}+\ldots+\frac{P_{-1}}{z-a}+\sum_{n=0}^{\infty} P_n(z-a)^n \quad (q\geqslant 1;\ P_{-q}\neq 0).$$

Преобразуем данную систему

$$\frac{dX}{dz}=PX, \tag{79}$$

положив

$$X=A(z)Y, \tag{80}$$

где $A(z)$ — матричная функция, регулярная в точке $z=0$ и принимающая в этой точке значение E:

$$A(z)=E+A_1(z-a)+A_2(z-a)^2+\ldots;$$

степенной ряд в правой части сходится при $|z-a|<r_1$.

Известный американский математик Г. Биркгофф в 1913 г. опубликовал теорему (см. [119]), согласно которой всегда можно подобрать преобразование (80) так, чтобы матрица коэффициентов преобразованной системы

$$\frac{dY}{dz}=P^*(z)Y \tag{79'}$$

содержала только отрицательные степени $z-a$:

$$P^*(z)=\frac{P^*_{-q}}{(z-a)^q}+\ldots+\frac{P^*_{-1}}{z-a}.$$

Теорема Биркгоффа вместе с полным ее доказательством приведена в книге Э. Л. Айнса «Обыкновенные дифференциальные уравнения»[1]). Там же на основе рассмотрения «канонических» систем (79′) проводится исследование поведения решения произвольной системы в окрестности особой точки.

---

[1]) См. [1], стр. 632—641. Биркгофф и Айнс формулируют теорему для особой точки $z=\infty$. Это не является каким-либо ограничением, поскольку любая особая точка $z=a$ преобразованием $z'=\frac{1}{z-a}$ может быть переведена в $z'=\infty$.

Между тем доказательство Биркгоффа содержит ошибку, а сама теорема неверна. В качестве опровергающего примера можно взять хотя бы пример, приведенный выше для опровержения утверждения Вольтерра[1]).

В этом примере $q=1$, $a=0$ и

$$P_{-1}=\begin{Vmatrix} 0 & 0 \\ 0 & -1 \end{Vmatrix}, \quad P_0=\begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}, \quad P_n=0 \text{ при } n=1,\ 2,\ \ldots$$

Применяя теорему Биркгоффа и подставляя в (79) вместо X произведение AY, мы после замены $\frac{dY}{dz}$ на $\frac{P^*_{-1}}{z}Y$ и сокращения на Y получим:

$$A\frac{P^*_{-1}}{z}+\frac{dA}{dz}=PA.$$

Приравнивая коэффициенты при $\frac{1}{z}$ и свободные члены, найдем:

$$P^*_{-1}=P_{-1},\ A_1P_{-1}-P_{-1}A_1+A_1'=P_0.$$

Полагая $A_1=\begin{Vmatrix} a & b \\ c & d \end{Vmatrix}$, получим:

$$\begin{Vmatrix} a & 0 \\ c & 0 \end{Vmatrix}-\begin{Vmatrix} 0 & 0 \\ -c & -d \end{Vmatrix}=\begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix}.$$

Это — противоречивое равенство.

В следующем параграфе мы выясним, к какому каноническому виду может быть преобразована система (79) при помощи преобразования (80) в случае регулярной особой точки.

§ 10. Регулярная особая точка

Исследуя поведение решения в окрестности особой точки, мы без нарушения общности рассуждения можем принять, что особой точкой является точка $z=0$[2]).

1. Пусть дана система

$$\frac{dX}{dz}=P(z)X, \tag{81}$$

где

$$P(z)=\frac{P_{-1}}{z}+\sum_{m=0}^{\infty}P_mz^m \tag{82}$$

и ряд $\sum\limits_{m=0}^{\infty}P_mz^m$ сходится внутри круга $|z|<r$.

---

[1]) В случае $q=1$ ошибочное утверждение Биркгоффа по существу совпадает с ошибкой Вольтерра (см. стр. 446).

[2]) Преобразованием $z'=z-a$ или $z'=\frac{1}{z}$ можно соответственно любую конечную точку $z=a$ или $z=\infty$ перевести в точку $z'=0$.

Положим

$$X = A(z)\,Y, \tag{83}$$

где

$$A(z) = E + A_1 z + A_2 z^2 + \ldots \tag{84}$$

Оставляя пока в стороне вопрос о сходимости ряда (84), постараемся так определить матричные коэффициенты A_m этого ряда, чтобы преобразованная система

$$\frac{dY}{dz} = P^*(z)\,Y, \tag{85}$$

где

$$P^*(z) = \frac{P^*_{-1}}{z} + \sum_{m=0}^{\infty} P^*_m z^m, \tag{86}$$

имела возможно более простой («канонический») вид[1]).

Подставляя в (81) вместо X произведение AY и используя (85), мы получим:

$$A(z)\,P^*(z)\,Y + \frac{dA}{dz}\,Y = P(z)\,A(z)\,Y.$$

Умножая обе части этого равенства справа на Y^{-1}, найдем:

$$P(z)\,A(z) - A(z)\,P^*(z) = \frac{dA}{dz}.$$

Заменяя здесь $P(z)$, $A(z)$, $P^*(z)$ рядами (82), (84), (86) и приравнивая в левой и правой частях равенства коэффициенты при одинаковых степенях z, получим бесконечную систему матричных уравнений для искомых коэффициентов A_1, A_2, ... [2]):

$$\left.\begin{aligned}
&1)\ P_{-1} = P^*_{-1},\\
&2)\ P_{-1}A_1 - A_1(P_{-1}+E) + P_0 = P^*_0,\\
&3)\ P_{-1}A_2 - A_2(P_{-1}+2E) + P_0A_1 - A_1P^*_0 + P_1 = P^*_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&(m+2)\ P_{-1}A_{m+1} - A_{m+1}[P_{-1}+(m+1)E] +\\
&\quad + P_0A_m - A_mP^*_0 + P_1A_{m-1} - A_{m-1}P^*_1 + \ldots + P_m = P^*_m.
\end{aligned}\right\} \tag{87}$$

2. Рассмотрим отдельно несколько случаев:

1° Матрица P_{-1} не имеет различных характеристических чисел, отличающихся друг от друга на целое число.

В этом случае при любом $k = 1, 2, 3, \ldots$ матрицы P_{-1} и $P_{-1} + kE$ не имеют общих характеристических чисел, и потому (см. гл. VIII, § 3)[3])

---

[1]) Мы будем стремиться к тому, чтобы в ряду (86) было конечное (и притом возможно меньшее) число коэффициентов P^*_m, отличных от нуля.

[2]) Во всех уравнениях, начиная со второго, мы заменим в силу первого уравнения матрицу P^*_{-1} на P_{-1}.

[3]) Можно, впрочем, доказать это и не опираясь на гл. VIII. Интересующее нас положение равносильно утверждению, что матричное уравнение

$$P_{-1}U = U(P_{-1} + kE) \tag{*}$$

имеет только нулевое решение $U = 0$. Поскольку матрицы P_{-1} и $P_{-1} + kE$ не имеют

матричное уравнение

$$P_{-1}U - U(P_{-1} + kE) = T$$

при любой правой части T имеет одно и только одно решение.

Это решение будем обозначать через

$$\Phi_k(P_{-1}, T).$$

Поэтому в уравнениях (87) можно положить все матрицы P_m^* $(m=0, 1, 2, \ldots)$ равными нулю и последовательно определить A_1, A_2. ... при помощи равенств

$$A_1 = \Phi_1(P_{-1}, -P_0), \quad A_2 = \Phi_2(P_{-1}, -P_1 - P_0A_1), \ \ldots$$

Тогда преобразованная система является системой Коши

$$\frac{dY}{dz} = \frac{P_{-1}}{z} Y,$$

и потому решение X исходной системы (81) имеет вид[1])

$$X = A(z) z^{P_{-1}}. \tag{88}$$

2° Среди различных характеристических чисел матрицы P_{-1} имеются числа, разность между которыми является целым числом; при этом матрица P_{-1} имеет простую структуру.

Обозначим через $\lambda_1, \lambda_2, \ldots, \lambda_n$ характеристические числа матрицы P_{-1}, расположенные так, чтобы имели место неравенства

$$\operatorname{Re}\lambda_1 \geqslant \operatorname{Re}\lambda_2 \geqslant \ldots \geqslant \operatorname{Re}\lambda_n. \tag{89}$$

Не нарушая общности рассуждений, мы можем матрицу P_{-1} заменить любой, ей подобной. Это следует из того, что, умножая обе части уравнения (81) слева на неособенную матрицу T, а справа — на T^{-1}, мы фактически заменяем все P_m на TP_mT^{-1} $(m = -1, 0, 1, 2, \ldots)$ (при этом и X заменяется на TXT^{-1}). Поэтому мы будем считать, что в рассматриваемом случае P_{-1} — диагональная матрица:

$$P_{-1} = \| \lambda_i \delta_{ik} \|_1^n. \tag{90}$$

Введем обозначения для элементов матриц P_m, P_m^* и A_m:

$$P_m = \| p_{ik}^{(m)} \|_1^n, \quad P_m^* = \| p_{ik}^{(m*)} \|_1^n, \quad A_m = \| x_{ik}^{(m)} \|_1^n. \tag{91}$$

Для определения A_1 мы воспользуемся вторым из уравнений (87). Это матричное уравнение можно заменить скалярными уравнениями

$$(\lambda_i - \lambda_k - 1)\, x_{ik}^{(1)} + p_{ik}^{(0)} = p_{ik}^{(0*)} \qquad (i, k = 1, 2, \ldots, n). \tag{92}$$

Если ни одна из разностей $\lambda_i - \lambda_k$ не равна единице, то мы можем положить $P_0^* = 0$. Тогда из (87_2)[2]) $A_1 = \Phi_1(P_{-1}, -P_0)$.

общих характеристических чисел, то существует такой многочлен $f(\lambda)$, для которого

$$f(P_{-1}) = 0, \quad f(P_{-1} + kE) = E.$$

Но из (*) следует:

$$f(P_{-1})\, U = U f(P_{-1} + kE).$$

Отсюда $U = 0$.

[1]) Формула (88) определяет одну интегральную матрицу системы (81). Произвольная интегральная матрица получается из (88) умножением справа на произвольную постоянную неособенную матрицу C.

[2]) Мы пользуемся обозначениями, введенными при разборе случая 1°.