

4.3.2. Линейные модели и линеаризация

Линеаризация является одним из простейших методов изучения нелинейных систем. Суть ее состоит в замене нелинейной системы на эквивалентную ей линейную. Линеаризованная модель не может полностью заменить нелинейную систему, но в некоторых отношениях ее поведение может быть (с определенным расхождением) идентичным поведению нелинейной системы. Поэтому существует возможность применить хорошо разработанные методы анализа линейных систем к линеаризованной модели, а полученные результаты интерпретировать для исходной нелинейной системы.

В технических науках, включая теорию управления, предполагается, что отклонения от некоторой рабочей точки малы, что позволяет нелинейные характеристики заменить эквивалентными линейными. В этом случае говорят о линеаризации вблизи рабочей точки, или о касательной аппроксимации. Линеаризованную модель далее исследуют при малых входных воздействиях.

Помимо касательной аппроксимации существуют и другие методы линеаризации, каждый из них для определенных задач обладает своими преимуществами. Например, метод гармонической линеаризации анализирует нелинейную систему не во временной области, а в частотной. Суть его в том, что рассматривается только основная составляющая выходного сигнала нелинейной системы, а высшие гармоники отбрасываются. Применение этого метода оправдано, когда нелинейная система имеет характеристику фильтра низких частот. Метод статистической линеаризации применяется во временной области. Эквивалентная линейная модель строится в предположении, что возмущения имеют гауссовское распределение. Существуют также комбинированные методы линеаризации.

В качестве примера далее рассмотрим только касательную аппроксимацию, причем когда коэффициенты линеаризованной модели получаются разложением в ряд.

В заключение приведем пример задачи, когда целесообразен переход от нелинейной модели к линеаризованной.

Пусть для некоторого объекта построена его математическая модель в виде системы НДУ, записанной в форме Коши:

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t)), \\ y(t) &= g(x(t), u(t)).\end{aligned}\tag{4.16}$$

В такой модели выделяют вектор состояния x , вектор входных воздействий u , вектор выходных (измеряемых, управляемых) координат y . Ясно, что на произвольном интервале времени T вид функций $x(t), y(t)$ будет зависеть от начальных условий и от вида функции на этом интервале. В таком случае задачей моделирования (или исследования) может быть построение функций $x(t), y(t)$ на интервале T , когда начальные условия и вид $u(t)$ на T заданы. Такая задача решается численным интегрированием системы НДУ, линеаризованная модель здесь является излишней.

Ясно, что вид $x(t), y(t)$ на интервале T будет зависеть от начальных условий и от вида $u(t)$ на T . Поэтому может быть поставлена следующая задача: выбрать такой вид $u_c(t)$, чтобы при заданных (или произвольных) начальных условиях и некотором (известном или неизвестном) виде $u_f(t)$ функция $x(t)$ (или $y(t)$) удовлетворяла бы на T заданным требованиям.

В этом случае целесообразно перейти к линейной модели, например, путем разложения в ряд Тейлора в окрестности статического режима. Переменные x, y, u_c, u_f заменяют их приращениями, а систему НДУ — системой ЛДУ с постоянными коэффициентами.

Если нелинейности в исходной системе НДУ достаточно гладкие, то нелинейный и линейный объекты имеют схожую динамику, что дает основание проводить исследования и синтез регуляторов для линейных систем и полученные результаты переносить на нелинейные, но уже с некоторой погрешностью. В большинстве случаев эта погрешность

оправдана широким набором готовых методов исследования линейных систем и синтеза линейных регуляторов.

В качестве примеров преимуществ линейных объектов перед нелинейными можно назвать следующие: выполнение принципа суперпозиции для входных воздействий; простота определения реакции линейной системы на типовые воздействия, что в совокупности с принципом суперпозиции позволяет использовать, например, методы спектрального анализа; наличие мощного математического аппарата для получения аналитических решений для линейных систем и др.

Вернемся к нелинейному объекту (4.16). Обозначив $z = (x, u)^T$, получим его запись в виде:

$$\begin{aligned}\dot{x}(t) &= f(z(t)), \\ y(t) &= g(z(t)).\end{aligned}\tag{4.17}$$

Левая часть первого уравнения (4.17) – функция $\dot{x}(t): R \rightarrow R^n$. Правая часть этого уравнения – сложная функция, т.е. композиция функций: $z(t): R \rightarrow R^{n+l}$ и $f(z(t)): R^{n+l} \rightarrow R^n$. Следует понимать, что функция $f(z)$ задана раз и навсегда, но вид функций $x(t)$ и $\dot{x}(t)$, удовлетворяющих уравнению на некотором интервале времени, зависит от значения $x(t_0)$ и вида $u(t)$ на этом интервале. Уравнение является дифференциальным и нелинейным, т.к. функция $f(z)$ в общем случае нелинейная. Аналогично рассматривается второе уравнение, но оно уже является алгебраическим и нелинейным.

Рассмотрим функцию $f(z): R^{n+l} \rightarrow R^n$. Предполагаем f непрерывной и дифференцируемой в некоторой области из R^{n+l} , содержащей любые допустимые значения вектора z . В этом случае можно построить ряд Тейлора функции f для любой точки z_0 из рассматриваемой области:

$$f(z) = f(z_0) + D_f(h) + o(h),$$

где вектор h вычисляется как $h = z - z_0$, а функция D_f есть линейная функция $D_f: R^{n+l} \rightarrow R^n$. Под $o(h)$ понимают некоторую функцию, значение которой на h пренебрежимо мало по сравнению со значением D_f на h . Смысл этого следующий: если требуется вычислить значение известной функции f в некоторой точке z , а ряд Тейлора задан, то с некоторой погрешностью значение $f(z)$ может быть вычислено как

$$f(z) \approx f(z_0) + D_f(h),$$

где z_0 известно, а $h = z - z_0$. При этом чем меньше значение (норма) h , тем качество приближения лучше. Отсюда следует выражение

$$f(z) - f(z_0) \approx D_f(h),$$

означающее, что изменение функции f относительно значения $f(z_0)$ при отклонении аргумента z от значения z_0 есть почти то же самое, что и значение функции D_f на $h = z - z_0$.

Линейная функция D_f называется дифференциалом функции f в точке z_0 . Функция D_f имеет вид:

$$D_f(h) = Jh,$$

где матрица J носит название матрицы Якоби функции f в точке z_0 .

Известно, что

$$J = \begin{pmatrix} \left. \frac{\partial f_1(z)}{\partial z_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_1(z)}{\partial z_{n+l}} \right|_{z=z_0} \\ \vdots & \dots & \vdots \\ \left. \frac{\partial f_n(z)}{\partial z_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_n(z)}{\partial z_{n+l}} \right|_{z=z_0} \end{pmatrix}.$$

В другой точке будет другая матрица, т.е. другой дифференциал. Теперь рассмотрим, как вышесказанное применимо к построению линейного аналога нелинейного уравнения. Пусть статический режим задан, т.е. найдено такое значение вектора z , при котором $\dot{x}(t) = 0 \forall t$. Выбираем это значение z в качестве z_0 .

Строим ряд Тейлора функции $f(z)$ в точке z_0

$$f(z) = f(z_0) + J(z - z_0) + o(z - z_0)$$

и отбрасываем последнее слагаемое. Полученное соотношение следует понимать так: если значения функции $z(t)$ в любой момент времени t не очень сильно отличаются от фиксированного значения z_0 , то значение сложной функции $f(z(t))$ в этот момент времени может быть приближенно получено как сумма вектора z_0 и произведения матрицы J на значение функции $h(t) = z(t) - z_0$ в этот момент времени. При этом матрица J фиксирована, т.е. значение $f(z)$ зависит от значения z , но не связано с моментом времени, в который z имеет это значение.

Функция $h(t)$ суть вектор-функция отклонений значений $z(t)$ от значения z_0 : $h(t) = (\Delta x(t), \Delta u(t))^T$. Итак, имеем:

$$\dot{x}(t) = f(z) \approx f(z_0) + J(\Delta x(t), \Delta u(t))^T$$

Далее отметим, что $f(z_0) = 0$, а также что $x(t) = x_0 + \Delta x(t)$, т.е. $\dot{x}(t) = \Delta \dot{x}(t)$.

Следовательно,

$$\Delta \dot{x}(t) = J(\Delta x(t), \Delta u(t))^T. \quad (4.18)$$

Полученное выражение есть система линейных дифференциальных уравнений, причем неоднородных (т.к. $\Delta u(t) \neq 0$).

Матрицу Якоби J разбивают на две части $J = [A \mid B]$:

$$A = \begin{pmatrix} \left. \frac{\partial f_1(z)}{\partial x_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_1(z)}{\partial x_n} \right|_{z=z_0} \\ \vdots & \dots & \vdots \\ \left. \frac{\partial f_n(z)}{\partial x_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_n(z)}{\partial x_n} \right|_{z=z_0} \end{pmatrix}, \quad B = \begin{pmatrix} \left. \frac{\partial f_1(z)}{\partial u_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_1(z)}{\partial u_l} \right|_{z=z_0} \\ \vdots & \dots & \vdots \\ \left. \frac{\partial f_n(z)}{\partial u_1} \right|_{z=z_0} & \dots & \left. \frac{\partial f_n(z)}{\partial u_l} \right|_{z=z_0} \end{pmatrix}$$

Эти матрицы называют матрицами состояния и входных воздействий соответственно. Причем, если входные воздействия разделены на

возмущающие и управляющие, то матрицу B также разбивают на матрицу возмущения B_1 и матрицу управления B_2 .

Для алгебраического уравнения поступают следующим образом: раскладывают $g(z)$ в ряд Тейлора в окрестности точки z_0

$$g(z) = g(z_0) + J_1(z - z_0) + o(z - z_0)$$

и рассматривают только $\Delta y(z) = J_1(z - z_0)$, называемое приращением выходных координат. Последнее соотношение представляет собой систему линейных алгебраических уравнений.

В результате получили:

$$\begin{aligned}\Delta \dot{x}(t) &= A \Delta x(t) + B u(t), \\ \Delta y(t) &= C \Delta x(t) + D u(t).\end{aligned}$$

Это и есть линеаризованный объект, то есть линейный аналог нелинейного объекта. Матрица C называется матрицей выхода, матрица D – матрицей обхода.

Часто вместо (4.16) используют упрощенные ММ, основанные на том, что процессы в системе протекают, мало отклоняясь от некоторой так называемой опорной траектории $\{\bar{x}(t), \bar{u}(t), \bar{y}(t)\}$, удовлетворяющей уравнениям

$$\begin{aligned}\dot{\bar{x}} &= F(\bar{x}, \bar{u}), \\ \bar{y} &= G(\bar{x}, \bar{u}).\end{aligned}\tag{4.19}$$

Тогда можно записать приближенную линеаризованную модель в отклонениях от этого режима:

$$\begin{aligned}\dot{\tilde{x}} &= A(t)\tilde{x} + B(t)\tilde{u}, \\ \tilde{y} &= C(t)\tilde{x} + D(t)\tilde{u},\end{aligned}\tag{4.20}$$

где $\tilde{x} = x - \bar{x}, \tilde{u} = u - \bar{u}, \tilde{y} = y - \bar{y}$,

$$\begin{aligned}A(t) &= \frac{\partial F}{\partial x}(\bar{x}(t), \bar{u}(t)), B(t) = \frac{\partial F}{\partial u}(\bar{x}(t), \bar{u}(t)), \\ C(t) &= \frac{\partial G}{\partial x}(\bar{x}(t), \bar{u}(t)), D(t) = \frac{\partial G}{\partial u}(\bar{x}(t), \bar{u}(t)),\end{aligned}$$

Если расчетный режим является установившимся, т.е. не зависит от времени, то коэффициенты в (4.20) также не зависят от времени: $A(t) \equiv A$, $B(t) \equiv B$ и т.д. Такие системы называются стационарными. Особенно часто на практике встречаются стационарные линейные непрерывные системы, описываемые более простыми уравнениями

$$\begin{aligned}\dot{x} &= Ax + Bu, \\ y &= Cx.\end{aligned}\tag{4.21}$$

Матрицы A, B, C являются параметрами модели (4.21). Если линеаризация приводит к большим погрешностям, то стараются, по возможности, выбрать математическую модель, линейную по параметрам

$$\dot{x} = A\psi(x, u),$$

где A – матрица параметров порядка $n \times N$; $\psi(\cdot)$ – нелинейная вектор-функция. К этому классу относятся, в частности, билинейные объекты, например $\dot{x} = a_1x + a_2xu + a_3u$, где $A = [a_1, a_2, a_3]$; $\psi(x, u) = (x, xu, u)^T$.

Сказанное выше относится и к уравнениям дискретных по времени систем. Уравнения дискретной системы в общем случае имеют вид

$$\begin{aligned}x_{k+1} &= F(x_k, u_k), \\ y_k &= G(x_k, u_k).\end{aligned}\tag{4.22}$$

Дискретным аналогом уравнений линейной стационарной системы (4.21) являются уравнения

$$\begin{aligned}x_{k+1} &= Px_k F + Qu_k, \\ y_k &= Rx_k.\end{aligned}\tag{4.23}$$

4.3.3 Дискретизация и континуализация

Дискретизация. Если исходное описание линейной системы непрерывно, можно перейти к дискретному описанию с помощью следующей процедуры.

Пусть состояние $x(t)$ системы (4.21) доступно измерению в дискретные моменты времени $t_k = kh$, $k = 0, 1, \dots$, где $h > 0$ – шаг дискретности. Пусть $u(t)$ постоянно на промежутках между моментами коррекции t_k . Тогда

динамику векторов $x_k = x(t_k)$ можно описать разностными уравнениями (3.23) в которых матрицы P и Q определяются соотношениями

$$\begin{aligned} P &= e^{Ah}, \\ Q &= A^{-1}(P - I_n)B. \end{aligned} \quad (4.24)$$

Здесь e^{Ah} – экспоненциал матрицы A , определяемый формулой

$$e^{Ah} = I_n + Ah + \frac{1}{2}A^2h^2 + \dots = \sum_{k=0}^{\infty} \frac{b^k A^k}{k!}. \quad (4.25)$$

Если предположение о кусочном постоянстве $u(t)$ не выполняется, то переход от (4.21) к (4.23) является приближенным, но его точность растет по мере уменьшения шага h , если скорость изменения входа (величина $\dot{u}(t)$) ограниченная. При достаточно малых h для вычисления e^{Ah} можно удерживать лишь первые несколько членов ряда (4.25) или аппроксимировать сумму (4.25) каким-либо способом.

Например, при переходе от (4.21) к (4.23) можно пользоваться формулой $e^{Ah} \approx I_n + Ah$, соответствующей численному интегрированию (4.21) методом Эйлера. При такой аппроксимации передаточные функции дискретной $W_d(z)$ и непрерывной $W_n(s)$ систем связаны соотношением

$$W_d(z) = W_n\left(\frac{1-z}{h}\right) \quad (4.26)$$

т.е. при переходе к дискретному времени в передаточной функции нужно

заменить s на $\frac{1-z^{-1}}{h}$. Целый ряд способов перехода от (4.21) к (4.24)

основан на аппроксимации матрицы e^{Ah} матричными дробями Паде (дробями, «числителем» и «знаменателем» которых являются матричные многочлены). Частными случаями этих способов является метод Тастина (формула Паде порядка (1,1)):

$$e^{Ah} \approx \left(I + \frac{Ah}{2}\right) \left(I - \frac{Ah}{2}\right)^{-1}, \quad (4.27)$$

приводящий к соотношению между передаточными функциями

$$W_d(z) = W_n \left(\frac{2}{h} \frac{1-z}{1+z} \right), \quad (4.28)$$

а также метод Дэвисона (формула Паде порядка (2,2)):

$$e^{Ah} \approx \left(I + \frac{Ah}{2} + \frac{A^2 h^2}{12} \right) \left(I - \frac{Ah}{2} - \frac{A^2 h^2}{12} \right)^{-1}. \quad (4.29)$$

Отметим, что формулы (4.27) и (4.29) дают устойчивые аппроксимации при $h > 0$ (разумеется, если A – гурвицева).

Заметим, что формула (4.24) для вычисления матрицы Q применима, если $\det A \neq 0$. Трудностей, связанных с вычислением Q при вырожденной матрице A , можно избежать, если при формальной подстановке выражения для $P = e^{Ah}$, полученного из аппроксимаций Тейлора или Паде в (4.24), произвести «сокращение» матрицы A . Тогда в выражение для Q матрица A^{-1} входить не будет. Например, аппроксимация по методу Эйлера $P = I + Ah$ приводит к формуле $Q = h \cdot B$.

Если непрерывная система нелинейная, то для перехода к ее дискретному описанию также можно использовать методы численного интегрирования. Например, метод Эйлера дает для системы (4.16) дискретное описание:

$$\begin{aligned} x_{k+1} &= x_k + hF(x_k, u_k, t_k), \\ y_k &= G(x_k, u_k, t_k). \end{aligned} \quad (4.30)$$

Континуализация – это переход от дискретной математической модели системы к непрерывной. Если дискретная модель системы имеет вид (4.23), то перейти к непрерывной модели (4.21) можно по формулам

$$A = \frac{1}{h} \ln P, \quad B = \frac{1}{h} \ln P (P - I)^{-1} Q, \quad (4.31)$$

вытекающим из (4.24), где $\ln P$ – логарифм матрицы, функция, обратная к экспоненциальной и также определяемая через ряд

$$\ln(I + X) = X - \frac{X^2}{2} + \dots + \frac{(-1)^{n+1}}{n} X^n, \quad (4.32)$$

сходящийся при $\|X\| < 1$ (здесь $X = P - I$). С точностью до величин порядка h^2 можно ограничиться формулами

$$A = \frac{1}{h}(P - I), B = \frac{1}{h}Q,$$

соответствующими методу Эйлера. Однако удобнее всего переходить от дискретной передаточной функции к непрерывной по формулам (4.26) и (4.28). Например, по методу Эйлера (4.26) достаточно заменить в передаточной функции $W_d(z^{-1})$ переменную z^{-1} на $1 - hp$.

4.3.4. Редукция моделей

При исследовании линейных систем получили распространение также методы упрощения описаний систем путем редукции (понижение порядка).

Возможность редукции математической модели можно определять по собственным числам матрицы состояния линейной или линеаризованной системы. Вещественные части собственных значений характеризуют скорость затухания переходных процессов. Если одно из собственных чисел минимум на порядок больше остальных, соответствующий ему переходный процесс закончится быстро и не окажет существенного влияния на переходный процесс модели в целом. В этом случае можно уменьшить порядок системы. Применительно к таким системам можно говорить о принципе подчинения.

В качестве примера можно рассмотреть нелинейную систему 2-го порядка

$$\begin{aligned}\dot{x}_1(t) &= \lambda_1 x_1 - x_1 x_2, \\ \dot{x}_2(t) &= -\lambda_2 x_2 + x_1^2,\end{aligned}$$

где коэффициент $\lambda_1 \geq 0$ и мал, а коэффициент $\lambda_2 \gg \lambda_1$

Если переменные x_1 и x_2 малы, то x_1 будет изменяться очень медленно. Поскольку $\lambda_2 \gg \lambda_1$, производной $\dot{x}_2(t)$ можно пренебречь в сравнении с величиной $\lambda_2 x_2$. Тогда исходную систему можно привести к виду

$$\dot{x}_1(t) = \lambda_1 x_1 - \frac{x_1^3}{\lambda_2}$$

где осуществлено алгебраическое исключение переменной x_2 . Поведение системы в основном определяется динамикой медленной подсистемы, которая как бы «управляет» быстрой подсистемой. При этом x_2 как бы подчинена x_1 . Медленная переменная x_1 в этом случае называется параметром порядка. В многомерных системах параметру порядка может быть подчинено весьма большое число других переменных. В общем случае параметров порядка может быть несколько, но часто это небольшое число, существенно меньшее размерности исходной системы.