**А.М. Инков**

# МОДЕЛИРОВАНИЕ И ИДЕНТИФИКАЦИЯ ОБЪЕКТОВ УПРАВЛЕНИЯ

## КОНСПЕКТ ЛЕКЦИЙ. ЛЕКЦИЯ № 3



Шымкент, 2011

## Тема 3 Статические и динамические модели в форме уравнения регрессии

### Лекция 3 Статические и динамические модели в форме уравнения регрессии (ч. 1)

В этой лекции мы кратко рассмотрим основные теоретические разделы теории вероятностей и линейной алгебры. Эти сведения понадобятся нам в следующих лекциях и на выполнении лабораторных и практических занятиях. Отметим, что использование таких систем, как Mathcad и MATLAB значительно упрощают практическое применение данного материала.

Регрессионный анализ является в настоящее время классическим статистическим методом. Благодаря своим широким возможностям различные регрессионные процедуры давно и успешно используются в инженерной практике для идентификации процессов, однако их применение к идентификации многомерных процессов в реальном масштабе времени стало возможным только с развитием и внедрением быстродействующих компьютеров**.**

Методы идентификации, основанные на регрессионных процедурах с использованием метода наименьших квадратов, применимы как к линейным, так и к нелинейным процессам и облегчают проведение идентификации по нескольким входам одновременно. Более того, регрессионные методы позволяют осуществлять идентификацию в реальном масштабе времени, поскольку они основаны на измерениях входных и выходных сигналов, которые можно получить в процессе нормального функционирования системы.

В течение периода, пока выполняются измерения для регрессионной идентификации, параметры идентифицируемого процесса принимаются стационарными или квазистационарными. Этот период должен быть не менее $mT$, где $T$ — интервал измерения,

**Элементы теории вероятностей. Основные понятия и определения.**

В теории вероятностей широко используются понятия события, вероятности, случайной величины.

Событие - всякий факт, который в результате опыта может произойти или не произойти.

Вероятность события есть численная мера степени объективной возможности этого события. Предположим, что рассматривается некоторый опыт или явление, в котором в зависимости от случая происходит или не происходит некоторое событие А.

Если условия опыта могут быть воспроизведены многократно, так что в принципе осуществима целая серия одинаковых и независимых друг от друга испытаний, то вероятность события А может быть вычислена по следующей формуле: $P(A) = m / n$, где n - общее число взаимно исключающих друг друга исходов; m – число исходов, которые приводят к наступлению события А.

Вероятность может принимать значения от 0 до 1. Событие, вероятность которого равна 0, называется невозможным, а событие, вероятность которого равна 1, называется достоверным. Несколько событий образуют полную группу событий, если в результате опыта должно непременно появиться хотя бы одно из них. Несколько событий называются несовместными в данном опыте, если никакие два из них не могут появиться вместе. Несколько событий называются равновозможными в данном опыте, если ни одно из этих событий не является объективно более возможным, чем другое. События называются независимыми, если появление одного из них не зависит от того, произошли ли другие события.

Случайной величиной называется величина, которая может принимать то или иное значение, неизвестное заранее.

Случайные величины могут быть двух типов:

- дискретные (прерывные), принимающие только отделенные друг от друга значения, которые можно пронумеровать;
- непрерывные (аналоговые), которые могут принимать любое значение из некоторого промежутка.

Иногда случайные величины, имеющие дискретную природу, рассматриваются как непрерывные. Такая замена оправдана в тех ситуациях, когда случайная величина принимает значения, которые незначительно отличаются друг от друга, так что замена дискретной случайной величины непрерывной практически не влияет на результаты расчетов. Например, время решения задачи в процессоре, определяемое как произведение среднего времени выполнения одной операции на число выполняемых операций, которое является дискретной случайной величиной, обычно рассматривается как непрерывная случайная величина, изменяющаяся в интервале от нуля до бесконечности.

Случайные величины часто обозначают большими буквами, а их возможные значения - соответствующими малыми буквами. Например, случайная величина X - число обращений к магнитным дискам в процессе решения задачи – может принимать значения x1=0, x2=1, x3=2, x4=3, ... .

Законы распределения случайных величин. Рассмотрим дискретную случайную величину X, принимающую значения x1, x2,...,xn. Величина X может принять каждое из этих значений с некоторой вероятностью. Обозначим через pi вероятность того, что случайная величина X примет значение xi: pi = Pr(X = xi) (i = 1, n). Если в результате опыта величина X принимает только одно из этих значений, то имеем полную группу несовместных событий и сумма вероятностей всех возможных значений случайной величины равна единице:

Эта суммарная вероятность каким-то образом распределена между отдельными значениями. Случайная величина будет полностью описана с вероятностной точки зрения, если мы зададим это распределение, т.е. установим, так называемый, закон распределения.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину говорят, что она подчинена данному закону распределения.

Дискретные законы распределения. Закон распределения дискретной случайной величины (дискретный закон распределения) может быть задан одним из следующих способов:

- аналитически в виде математического выражения, отражающего зависимость вероятности от значения случайной величины;
- таблично в виде ряда распределения случайной величины, в котором перечислены возможные значения случайной величины и соответствующие им вероятности;
- графически в виде многоугольника распределения, при котором по оси абсцисс откладываются возможные значения случайной величины, а по оси ординат - вероятности этих значений.

Числовые характеристики случайных величин. Числовые характеристики позволяют выразить в сжатой форме наиболее существенные особенности распределения случайной величины. Во многих случаях достаточно бывает указать только отдельные числовые параметры, до некоторой степени характеризующие существенные черты распределения случайной величины: например, какое-то среднее значение, около которого группируются возможные значения случайной величины; какое-либо число, характеризующее степень разбросанности этих значений относительно среднего; число, характеризующее асимметрию (или “скошенность”) распределения; число, характеризующее “крутость”, то есть островершинность или плосковершинность распределения и так далее.

В теории вероятностей используется большое количество числовых характеристик, имеющих различное назначение и различные области применения. Из них наиболее часто используются начальные и центральные моменты различных порядков, каждый из которых описывает то или иное свойство распределения. Начальные моменты рассматриваются относительно начала координат, а центральные моменты – относительно математического ожидания, то есть центра распределения.

Применение числовых характеристик существенно облегчает решение многих вероятностных задач. Особое значение это имеет при решении сложных задач, когда использование законов распределений приводит к громоздким выкладкам и не позволяет получить результаты в явном виде. Очень часто удается решить задачу до конца, оставляя в стороне законы распределения и оперируя одними числовыми характеристиками. Если в задаче фигурирует большое количество случайных величин, то в этих случаях по существу задачи для исчерпывающего суждения о результирующем законе распределения не требуется знать законов распределения отдельных случайных величин, фигурирующих в задаче; достаточно знать лишь некоторые числовые характеристики этих величин.

Пусть дискретная случайная величина X может принимать значения x1, x2,…,xk. Тогда: Частота появления события X= xi (частота = min) – отношение числа опытов mi в которых X приняла значение xi к общему числу опытов n.

Вероятность события X=xi [обозначается P(X=xi)] стремится к min при большом n: pi= P(X=xi) ≈ min

**Аксиомы теории вероятностей Колмогорова.**

1)Вероятность появления случайного события A – неотрицательное число: P(A)≥ 0

2)Вероятность достоверного события U – равна единице: P(U)=1; а невозможного события V – равна нулю: P(V)=0, т.е. $0 \le P \le 1$

3)Вероятность того, что наступит хотя бы одно из нескольких несовместимых событий A1 + A2 + … + An равно сумме вероятностей этих событий (теорема сложения вероятностей): P(A1 + A2 + … + An) = P(A1) + P(A2) + … + P(An)

Сумма событий (A1 + A2 + … + An) это событие, соответствующее появлению хотя бы одного из событий Ai, а произведение, это событие, соответствующее появлению всех событий Ai. Сумма вероятностей всех возможных значений случайной величины равно единице.

События независимые, если вероятность любого из них не зависит от того, произойдет или нет любое из остальных. Вероятность произведения нескольких независимых событий равна произведению вероятностей этих событий:

P(A1 · A2 · … · An) = P(A1) · P(A2) · … · P(An).

Событие A зависимо от события B, если вероятность события A меняется в зависимости появления события B. Вероятность события A, вычисленная при условии, что произошло события B – условная вероятность события A, обозначается P(A/B).

Для зависимых событий вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое произошло. P(AB) = P(A) · P(B/A). Аналогично, если событие B предшествует событию A и влияет на него, то P(AB) = P(B) · P(A/B).

Пример. Требуется определить надежность системы автоматического управления, состоящей из трех последовательно соединенных элементов, каждый из которых может выйти из строя. Вероятности безотказной работы каждого из них соответственно равны: P(A1) = 0,9; P(A2) = 0,85; P(A3) = 0,82. В соответствии с теоремой умножения P(A1 · A2 · A3) = P(A1) · P(A2) · P(A3) = 0,9 ·0,85 ·0,82 = 0,6273

Математическое ожидание и дисперсия, их оценка и свойства. Оценка математического ожидания переменной X (обозначается mX или $\overline{X}$ ):

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (3.1)$$

n – количество опытов

Оценка дисперсии переменной X (обозначается $\sigma_X^2$ или $S_X^2$ ):

$$S_X^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{f} \qquad (3.2)$$

Число степеней свободы f = [общее число измерений] – [число оценок, уже рассчитанным по этим измерениям и примененным в текущей формуле]. В данном случае уже рассчитана и используется величина $\overline{X}$, т.е. $f = n - 1$.

$S_X = \sqrt{S_X^2}$, среднее квадратичное отклонение (ошибка, стандарт).

Оценка дисперсии воспроизводимости Y (обозначается $\sigma_Y^2$ или $S_Y^2$):

$$S_{BOCПР}^2 = \frac{\sum_{i=1}^{n_0}(Y_i - \overline{Y})^2}{n_0 - 1} \qquad (3.\ 3)$$

$$\text{или } S_{BOCПР}^2 = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n_0}(Y_{ji} - \overline{Y}_j)^2}{m \cdot (n_0 - 1)} \qquad (3.3А)$$

$S_{BOCП} = \sqrt{S_{BOCП}^2}$ - ошибка опыта (ошибка воспроизводимости, среднеквадратическая ошибка, среднеквадратическое отклонение -СКО).(См [1] стр.37)

Оценка остаточной дисперсии Y (дисперсии адекватности):

$$S_{OCT}^2 = \frac{\sum_{i=1}^{n_0}(Y_i - \widehat{Y}_i)^2}{n_0 - L} \qquad (3.4)$$

$\widehat{Y}$ - расчетное значение выхода; L – количество коэффициентов в уравнении регрессии;

Нормальное распределение.

$$f(x) = \frac{1}{S_X \cdot \sqrt{2\pi}} \cdot e^{-\frac{(X - \overline{X})^2}{2 \cdot S_X^2}} \quad (-\infty < x < \infty) \qquad (3.5)$$

**Критерий Стьюдента.** Критерий позволяет определить значимость коэффициентов уравнения регрессии. $b_i$. (используется для проверки гипотезы значимости коэффициентов). Для этого для каждого из L коэффициентов $b_i$. Рассчитывают по формуле:

$$t_{b_i} = \frac{|b_i|}{S_{b_i}}, \qquad (3.6)$$

если $t_{b_i} > t_{TAБ}(f)$, то $b_i$ считается значимым, в противном случае он не значимым и приравнивается к нулю.

$$S_{b_i} = \sqrt{\frac{S^2_{BOCПР}}{k}} = \frac{S_{BOCПР}}{\sqrt{k}}, \; S_{BOCПР} = \sqrt{S^2_{BOCПР}} \qquad (3.7)$$

$S_{b_i}$ - среднеквадратическая ошибка в определении коэффициента регрессии $b_i$;

$t_{TAБ}(f)$ - табличное значение критерия Стьюдента при f = n0 – 1 или f =m· (n0 – 1);

Иногда незначимость $b_i$ может быть вызвана и другими причинами, например, неверным интервалом варьирования при подготовке к эксперименту. Ниже приведен пример на Mathcad использования критерия Стьюдента. [1, стр.164].

Сделаем замечание относительно проблемы оценки значимости коэффициентов уравнения регрессии (2.2). Если какой либо коэффициент незначим, то он может быть принят равным нулю, т.к. его влияние на результат расчета по формуле (2.2) мало. Чтобы оценить допустимый предел требуемой точности определения расчетного значения выхода $\hat{Y}$ надо учесть, что в инженерной и научной практике обычно достаточно 3-5 значащих цифр при записи и использовании числовых расчетов. Для обоснованного отброса незначимых коэффициентов необходимо иметь информацию о дисперсии, присущей экспериментально найденным значениям выхода *Y*, а также значениям расчетным $\hat{Y}$. Строго говоря, экспериментально найденные значения Y являются случайными величинами. Дисперсия случайной величины это мера разброса данной случайной величины, то есть мера её отклонения от математического ожидания (т.е. среднеарифметического от ряда параллельных экспериментальных измерений). Если дисперсия эксперимента превышает допустимую величину, то доверять таким экспериментам нельзя, т.к. они невоспроизводимы. Это может происходить при низкой точности измерительных приборов, неправильной методики эксперимента, неучета дополнительных факторов, влияющих на результаты эксперимента, взаимного влияния входов друг на друга и т.п. Вообще говоря, отбрасывание незначимых коэффициентом актуально, когда уравнение (1) используется при ручном расчете, например с использованием калькуляторов. При использовании же при расчете компьютерных технологий отбрасывание незначимых коэффициентов можно не проводить.

**Оценка значимости коэффициентов уравнения регрессии по критерию Стьюдента.**

Исходные коэффициенты $b^T = (8.5 \quad 2.5 \quad -0.5 \quad 3.5 \quad -0.5 \quad 0.5 \quad -1.5 \quad -0.5)$

Проведено 3 паралельных опыта $Y0_1 = 8 \quad Y0_2 = 9 \quad Y0_3 = 8.8$

$n0 = 3 \quad i = 1 .. n0$

$L = 8 \quad j = 1 .. L$

Матожидание Y0= $Y0cp = \dfrac{\left(\sum_{i=1}^{n0} Y0_i\right)}{n0} \quad Y0cp = 8.6$

Определение дисперсии воспроизводимости Sb

$$S2b = \frac{\left[\sum_{m=1}^{M} \left(Y0_m - y0\right)^2\right]}{f2} \qquad S = \sqrt{S2b} \quad Sb = \frac{S}{\sqrt{N}}$$

$S2b = 0.28 \qquad S = 0.529 \quad Sb = 0.187$

$Stud = \begin{pmatrix} 12.71 \\ 4.30 \\ 3.18 \\ 2.78 \\ 2.57 \\ 2.45 \\ 2.37 \\ 2.31 \\ 2.26 \\ 2.23 \end{pmatrix}$ <---при-f=2

$T_j = \dfrac{|b_j|}{Sb} \quad T^T = (45.434 \quad 13.363 \quad 2.673 \quad 18.708 \quad 2.673 \quad 2.673 \quad 8.018 \quad 2.673)$

Число степеней свободы дисперсии воспроизводимости= $f = n0 - 1 \quad f = 2$

$b_j = if\left[T_j < Stud_{(f)}, 0, b_j\right]$

Значимые коэффициенты $b^T = (8.5 \quad 2.5 \quad 0 \quad 3.5 \quad 0 \quad 0 \quad -1.5 \quad 0)$

Таким образом, в этом примере Lзн = 4

**Критерий Кохрена.** Применяется для проверки воспроизводимости опытов (для проверки гипотезы воспроизводимости опытов):

$$G_P = \frac{S^2_{MAX}}{\sum_{i=1}^{n_0} S_i^2}, \qquad (3.8)$$

т.е. расчетное значение GP определяется как отношение наибольшей из оценок дисперсий к сумме всех найденных оценок дисперсий.

Если расчетное GP > GTAБ, то дисперсии неоднородны, т.е. значения Y не подчиняются нормальному закону распределения, а опыты невоспроизводимы). $S^2_{MAX}$ - наибольшая из оценок выборочных дисперсий;

n0 – общее число сравниваемых дисперсий (количество параллельных (дублирующих) опытов);

$\sum_{i=1}^{n_0} S_i^2$ - сумма всех оценок дисперсий. Количество опытов в серии должно быть одинаковым, в противном случае можно воспользоваться критерием Бартлетта. Необходимо знать: n0 и число степеней свободы f = n0 - 1. Если опыты невоспроизводимы, то можно попытаться выявить и устранить источники невоспроизводимости (помехи), увеличить точность измерения и т.п. Если воспроизводимость не может быть обеспечена, то и результаты эксперимента не могут быть использованы для дальнейшей математической обработки. Например, (см. таб.) рассмотрим эксперимент, состоящий из трех серий (m = 3) по два параллельных опыта (n0 = 2). В них выход Y зависит от двух факторов X1 и X2.

| номера серий опытов | Параллельные опыты | | | | $Y_i$2средн% | Дисперсия Y= $S_i$^2 |
|---|---|---|---|---|---|---|
| | Условия опытов | | Выходы | | | |
| | $X_1$ гр.С | $X_2$ % | $Y_i$1 % | $Y_i$2 % | | |
| 1 | 24 | 45 | 35 | 36 | 35,5 | 0,5 |
| 2 | 24 | 55 | 39,3 | 38,1 | 38,7 | 0,72 |
| 3 | 25 | 45 | 31,8 | 33,4 | 32,6 | 1,28 |
| | | | | | Сумма = | 2,5 |

Расчетное значение критерия Кохрена GP = 1.28/2.5 = 0.51, в таб. (при m = 3 и f = n0 – 1 = 1) находим GTAБ = 0.967, т.к. GP < GTAБ, то опыты воспроизводимы, а оценки дисперсий можно считать однородными.

Вычислим также оценку дисперсии воспроизводимости:

$$S^2_{BOCПР} = \frac{\sum_{i=1}^{n_0} (Y_i - \bar{Y})^2}{n_0 - 1} = (0.50 + 0.72 + 1.28)/3 = 0.83,$$ с ней связано число степеней свободы f = m·(n0 - 1) = 3·(2-1) = 3.

**Критерий Фишера.** Используется для проверки адекватности уравнения регрессии.

Расчетное значение критерия Фишера определяют как:

$$F_{PAC} = \frac{S^2_{OCT}}{S^2_{ВОСПР}} \qquad (3.9)$$

Оценка остаточной дисперсии (дисперсии адекватности) рассчитывают по формуле: $S^2_{OCT} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n - L_{ЗН}}$ $\hat{Y}$ - значения Y, рассчитанные по уравнению регрессии на основе значимых коэффициентов. Если выполняется условие $F_{PAC} \le F_{ТАБ}$, уравнение регрессии адекватно описывает экспериментальные данные. $F_{ТАБ}$ определяется при известных значениях степеней свободы f1 для числителя и f2 для знаменателя.

f1 = n – L3H и f2 = n0 – 1.

Обычно используют уровень значимости p=0,05.

**Выборочный коэффициент корреляции**[См.1, стр.121]**:**

$$r = \frac{\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right)\cdot\left(Y_i - \overline{Y}\right)\right]^2}{\sqrt{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \cdot \sum_{i}^{n}\left(Y_i - \overline{Y}\right)^2 \cdot\right]}} \text{ или } r^*_{XY} = \frac{\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right)\cdot\left(Y_i - \overline{Y}\right)\right]^2}{(n-1)\cdot S_X \cdot S_Y}$$

Рассмотрим пример на Mathcad, для данных приведенный в [1, стр.164].

**Проверка адекватности уравнения регрессии по критерию Фишера.**

Значимые коэффициенты $b^T = (8.5 \quad 2.5 \quad 0 \quad 3.5 \quad 0 \quad 0 \quad -1.5 \quad 0)$

$$Yr_i = b_1 \cdot X_{i,1} + b_2 \cdot X_{i,2} + b_3 \cdot X_{i,3} + b_4 \cdot X_{i,4} + b_5 \cdot X_{i,5} + b_6 \cdot X_{i,6} + b_7 \cdot X_{i,7} + b_8 \cdot X_{i,8}$$

Число степеней свободы дисперсии воспроизводимости= $f2 = n0 - 1 \quad f2 = 2$

$$S2o = \frac{\left[\sum_{i=1}^{N} \left(Y_i - Yr_i\right)^2\right]}{f1}$$

Число степеней свободы остаточной дисперсии f1:

$f1 = N - 4 \qquad f1 = 4$

Остаточная дисперсия= $S2o = 2$

$$F = \frac{S2o}{S2b}$$

$$Fisher = \begin{pmatrix} 164.4 & 199.5 & 215.7 & 224.6 & 230.2 \\ 18.50 & 19.20 & 19.20 & 19.30 & 19.30 \\ 10.10 & 9.600 & 9.300 & 9.100 & 9.000 \\ 7.700 & 6.900 & 6.600 & 6.400 & 6.300 \\ 6.600 & 5.800 & 5.400 & 5.200 & 5.100 \end{pmatrix}$$

$$Yr = \begin{pmatrix} 1 \\ 6 \\ 4 \\ 9 \\ 11 \\ 16 \\ 8 \\ 13 \end{pmatrix} \qquad Y = \begin{pmatrix} 2 \\ 6 \\ 4 \\ 8 \\ 10 \\ 18 \\ 8 \\ 12 \end{pmatrix}$$

$Fisher_{f2,f1} = 19.3 \qquad F = 7.143$

Так как F<Fisher , то уравнение адекватно описывает эксперимент

**Основы матричной алгебры (операции с матрицами)**

Матрица - математический объект, записываемый в виде прямоугольной таблицы чисел (или элементов кольца) и допускающий алгебраические операции (сложение, вычитание, умножение) между ним и другими подобными объектами. Обычно матрицы представляются двумерными (прямоугольными) таблицам. Использование матричного аппарата позволяет значительно упростить вычислительные операции с использованием линейных математических моделей.

У каждого элемента матрицы есть два нижних индекса ($a_{ij}$) — первый «i» обозначает номер строки, в которой находится элемент, а второй «j» — номер столбца.

Говорят «матрица размера $m \times n$,» подразумевая, что в матрице m строк и n столбцов. В такой матрице индексы элементов удовлетворяют неравенствам $0 < i \leqslant m,$ (если индексы отсчитываются от единицы) или $0 \leqslant i < m,$ (если индексы отсчитываются от нуля).

Операции над матрицами. Пусть $a_{ij}$ - элементы матрицы A, а $b_{ij}$ - элементы матрицы B.

Линейные операции (это умножение на число, сложение и вычитание).

Умножение матрицы A на число λ (обозначение: λA) заключается в построении матрицы B, элементы которой получены путём умножения каждого элемента матрицы A на это число, то есть каждый элемент матрицы B равен: $b_{ij} = \lambda a_{ij}$

Сложение матриц A + B есть операция нахождения матрицы C, все элементы которой равны попарной сумме всех соответствующих элементов матриц A и B, то есть каждый элемент матрицы C равен: $c_{ij} = a_{ij} + b_{ij}$

Вычитание матриц A − B определяется аналогично сложению, это операция нахождения матрицы C, элементы которой: $c_{ij} = a_{ij} - b_{ij}$

Примеры:

$$A+B = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 1 & 0 \\ 8 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2+3 & 0+1 & -1+0 \\ 1+8 & 3+2 & 0+3 \end{pmatrix} = \begin{pmatrix} 5 & 1 & -1 \\ 9 & 5 & 3 \end{pmatrix}$$

$$A-B = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \end{pmatrix} - \begin{pmatrix} 3 & 1 & 0 \\ 8 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2-3 & 0-1 & -1-0 \\ 1-8 & 3-2 & 0-3 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 \\ -7 & 1 & -3 \end{pmatrix}$$

Нелинейные операции (это умножение матриц, возведение в степень, транспонирование, обращение). Умножение матриц (обозначение: AB, реже со знаком умножения $A \times B$) — есть операция вычисления матрицы C, элементы которой равны сумме произведений элементов в соответствующей строке первого множителя и столбце второго, а именно:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

В первом множителе должно быть столько же столбцов, сколько строк во втором. Если матрица A имеет размерность $m \times n$, B - $n \times k$, то размерность их произведения AB = C есть $m \times k$. Схема этой процедуры показана на рисунке 3.1

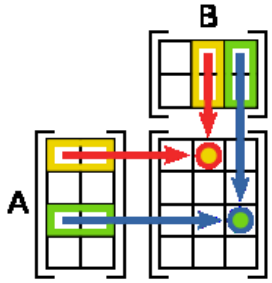



Рисунок 3.1 – Схема умножения матриц

Примеры умножения матриц:

$$FL = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \times \begin{pmatrix} g & i & k \\ h & j & l \end{pmatrix} = \begin{pmatrix} (a \cdot g + d \cdot h) & (a \cdot i + d \cdot j) & (a \cdot k + d \cdot l) \\ (b \cdot g + e \cdot h) & (b \cdot i + e \cdot j) & (b \cdot k + e \cdot l) \\ (c \cdot g + f \cdot h) & (c \cdot i + f \cdot j) & (c \cdot k + f \cdot l) \end{pmatrix}$$

$$AB = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \times \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-1) + 3 \cdot (-2) & 2 \cdot 2 + 3 \cdot 3 \\ 5 \cdot (-1) + 7 \cdot (-2) & 5 \cdot 2 + 7 \cdot 3 \end{pmatrix} = \begin{pmatrix} -8 & 13 \\ -19 & 31 \end{pmatrix}$$

$$BA = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} \times \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} -1 \cdot 2 + 2 \cdot 5 & -1 \cdot 3 + 2 \cdot 7 \\ -2 \cdot 2 + 3 \cdot 5 & -2 \cdot 3 + 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} 8 & 11 \\ 11 & 15 \end{pmatrix}$$

$$A := \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix} \quad B := \begin{pmatrix} 5 & 4 & -3 \\ 0 & 9 & 6 \\ 5 & 8 & 3 \end{pmatrix} \quad A \cdot B = \begin{pmatrix} 20 & 46 & 18 \\ -5 & 41 & 33 \\ 20 & 46 & 18 \end{pmatrix} \quad B \cdot A = \begin{pmatrix} -2 & 24 & 6 \\ -3 & 57 & 18 \\ 0 & 56 & 24 \end{pmatrix}$$

Возводить в степень можно только квадратные матрицы, например:

$$A := \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix} \quad A^3 = \begin{pmatrix} -4 & 118 & 42 \\ -32 & 97 & -27 \\ -4 & 118 & 42 \end{pmatrix} \quad A \cdot A \cdot A = \begin{pmatrix} -4 & 118 & 42 \\ -32 & 97 & -27 \\ -4 & 118 & 42 \end{pmatrix}$$

Транспонирование матрицы (обозначение: $A^T$) — операция, при которой матрица отражается относительно главной диагонали, то есть: $a_{ij}^T = a_{ji}$

Если A — матрица размера $m \times n$, то $A^T$ — матрица размера$n \times m$.

Свойства операций над матрицами

- ассоциативность сложения: A + (B + C) = (A + B) + C.
- коммутативность сложения: A + B = B + A.
- ассоциативность умножения: A(BC) = (AB)C.

Вообще говоря, умножение матриц не коммутативно: $AB \neq BA$

Дистрибутивность умножения относительно сложения:

A(B + C) = AB + AC; (B + C)A = BA + CA.

Свойства операции транспонирования матриц:

$(A^T)^T = A$; $(AB)^T = B^T A^T$: $(A^{-1})^T = (AT)^{-1}$, (если обратная матрица $A^{-1}$ существует); $(A+ B)^T = A^T + B^T$; $detA = detA^T$

Если количество строк матрицы равно количеству столбцов, то такая матрица называется квадратной. Для квадратных матриц существует единичная матрица E (аналог единицы для операции умножения чисел) такая, что умножение любой матрицы на неё не влияет на результат, а именно: EA = AE = A. У единичной матрицы единицы стоят только по главной диагонали, остальные элементы равны нулю.

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Для некоторых квадратных матриц можно найти так называемую обратную матрицу. Обратная матрица $A^{-1}$ такова, что если умножить матрицу на неё, то получится единичная матрица: $AA^{-1} = E$. Обратная матрица существует не всегда. Матрицы, для которых обратная существует, называются невырожденными (или регулярными), а для которых нет — вырожденными (или **сингулярными**). Матрица невырождена, если все ее строки (столбцы) линейно независимы как векторы. Максимальное число линейно независимых строк (столбцов) называется рангом матрицы. Определителем (детерминантом) матрицы называется значение нормированной кососимметрической (антисимметрической) полилинейной формы валентности $(p;\ 0)$ на столбцах матрицы. Квадратная матрица над числовым полем вырождена тогда и только тогда, когда ее определитель равен нулю. У сингулярной матрицы определитель равен нулю или близок к нулю. Обычно это происходит, когда элементы матрицы значительно (на много порядков) отличаются друг от друга.

**Сингулярность** матрицы иногда делает невозможность выполнять решение задчи регрессионного анализа и решения уравнений.

Матрица как запись коэффициентов системы линейных уравнений. Систему из m уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

можно представить в матричном виде:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

и тогда всю систему можно записать так: AX = B, где A имеет смысл таблицы коэффициентов $a_{ij}$ системы уравнений. Если m = n и матрица A **невырожденная**, то решение этого уравнения состоит в нахождении обратной матрицы $A^{-1}$, поскольку умножив обе части уравнения на эту матрицу слева: $A^{-1}AX = A^{-1}B$ $A^{-1}A$ — превращается в E (единичную матрицу). И это даёт возможность получить столбец корней уравнений: $X = A^{-1}B$.

**Векторы**. Нами будут использоваться понятие вектор в смысле - одномерная матрица, т.е. вектор столбец. Для удобства записи вектора в строчку иногда применяют запись вида $\mathbf{B^T}$**=** вместо записи **B=**. Для вектора справедливы все приведенные выше правила операций с матрицами, например:

$$X := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad Y := \begin{pmatrix} 5 \\ 4 \\ 7 \end{pmatrix} \quad X \cdot Y = 34 \quad Y \cdot X = 34 \quad X \cdot Y^T = \begin{pmatrix} 5 & 4 & 7 \\ 10 & 8 & 14 \\ 15 & 12 & 21 \end{pmatrix} \quad X^T \cdot Y = 34$$

**Основная литература**

1. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
2. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.

**Дополнительная литература**

3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
4. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.