# Лекция 19

# СТАТИСТИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ СОСТОЯНИЙ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

Основная задача технической диагностики — на основе анализа диагностических данных определить состояние функционирующего технического объекта. Определение состояния объекта проводится, как правило, в условиях ограниченной информации, поэтому не исключен риск ошибки при постановке диагноза. Ошибочные решения влекут за собой неминуемые потери. Они могут быть существенными и даже катастрофическими, если ошибки послужили причиной тяжелых техногенных аварий.

В таких условиях специалистам, делающим заключение о техническом состоянии контролируемого объекта, необходимо иметь четкие алгоритмы и правила принятия решения. В них должны быть учтены как неизбежность ошибочных решений, так и тяжесть их последствий.

С другой стороны, с помощью этих правил можно оценить эффективность используемых методов и средств диагностики. Если результаты анализа покажут неприемлемую долю ошибочных решений или связанные с ними большие потери, то это послужит основанием для разработки новых или совершенствования существующих методов диагностического обслуживания.

Для обоснованной постановки диагнозов с учетом перечисленных факторов применяются *методы распознавания* и *статистические методы принятия решений*. Они были разработаны в области технической кибернетики, теории информации, теории связи и радиолокации и в дальнейшем нашли применение в технической диагностике. Этим объясняется характерная для указанных наук терминология (см. ниже), перекочевавшая в сферу диагностики.

Методы распознавания и статистические методы принятия решений относятся к так называемым *параметрическим* методам постановки диагнозов. Чтобы воспользоваться ими, необходимо знать плотности вероятности диагностических параметров, если они являются непрерывными случайными величинами, или вероятности наличия признаков для различных состояний диагностируемого объекта.

Основы применения методов распознавания и статистических методов принятия решений в технической диагностике изложены И.А. Биргером, [1] (здесь и далее в лекциях 19-24 ссылки даны согласно списку литературы по данному разделу курса, приведенному в конце лекции 24). Более полное описание этих методов можно найти в специальных работах

[2, 3]. В нашем курсе мы сделаем упор на практическое применение указанных методов, снабдив материал примерами и подробным решением задач из области диагностики ядерных энергетических установок.

## Метод Байеса

Метод Байеса принятия решений о состоянии диагностируемых объектов основан на использовании обобщенной формулы Байеса. Для ее применения требуется большой объем предварительной информации, которая может быть накоплена в процессе эксплуатации объектов или в результате специально организованных экспериментальных исследований. При наличии необходимых данных с помощью метода Байеса можно оценить состояние объектов с высоким уровнем распознавания (см. ниже), поэтому его относят к эффективным методам принятия решений.

Будем считать, что диагностируемый объект может находиться в одном из возможных состояний $D_i$ $\left(i = 0, 1, \ldots, N\right)$ и характеризуется сложным признаком $\mathbf{k} = \left(k_1, k_2, \ldots, k_m\right)$. В соответствии с методом Байеса рассчитываются условные вероятности $P\left(D_i/\mathbf{k}\right) = P\left(D_i/k_1, k_2, \ldots, k_m\right)$. Это — *апостериорные*, то есть установленные после диагностирования, вероятности состояний $D_i$ при условии, что объект обладает признаком $\mathbf{k}$.

*Решающее правило* — правило принятия решения о нахождении объекта в состоянии $D_i$ — состоит в том, что объект с комплексным признаком $\mathbf{k}$ относят к указанному состоянию, если апостериорная вероятность этого состояния максимальна

$$P\left(D_i/\mathbf{k}\right) > P\left(D_j/\mathbf{k}\right), \qquad j = 0, 1, \ldots, N;\ i \neq j. \tag{Л19.1}$$

Это правило усиливают введением *уровня распознавания* $P^*$, обычно $P^* \geq 0,9$. Решение о диагнозе $D_i$ принимается, если выполнены условия (Л19.1) и

$$P\left(D_i/\mathbf{k}\right) \geq P^*. \tag{Л19.2}$$

При этом вероятность альтернативных диагнозов не превышает $1 - P^*$. В ином случае отказываются от постановки диагноза, для постановки которого требуется дополнительная информация.

Рассмотрим процедуру определения апостериорных вероятностей $P\left(D_i/\mathbf{k}\right)$. Вероятность двух событий — нахождения объекта в состоянии $D_i$ и наличия у него сложного признака $\mathbf{k}$ — определяется формулами теории вероятностей

$$P\left(D_i, \mathbf{k}\right) = P\left(D_i\right)P\left(\mathbf{k}/D_i\right) = P\left(\mathbf{k}\right)P\left(D_i/\mathbf{k}\right), \qquad \text{(Л19.3)}$$

где $P\left(D_i\right)$ — априорная (известная до диагностирования по результатам специально проведенных предварительных исследований) вероятность состояния $D_i$; $P\left(\mathbf{k}/D_i\right)$ — условная вероятность наличия у объекта признака $\mathbf{k}$, если он находится в состоянии $D_i$, $P\left(\mathbf{k}\right)$ — вероятность наличия у объекта комплексного признака $\mathbf{k}$ независимо от его состояния. Поскольку объект при любом $\mathbf{k}$ может находиться только в одном состоянии, то

$$\sum_{i=0}^{N} P\left(D_i/\mathbf{k}\right) = 1. \qquad \text{(Л19.4)}$$

Просуммировав обе части правого равенства в (Л19.3) по $i$, определим вероятность наличия у объекта признака $\mathbf{k}$:

$$P\left(\mathbf{k}\right) = \sum_{i=0}^{N} P\left(D_i\right)P\left(\mathbf{k}/D_i\right). \qquad \text{(Л19.5)}$$

С учетом этого равенства из (Л19.3) следует *обобщенная формула Байеса* для вычисления апостериорных вероятностей диагнозов $D_i$:

$$P\left(D_i/\mathbf{k}\right) = \frac{P\left(D_i\right)P\left(\mathbf{k}/D_i\right)}{\sum_{j=0}^{N} P\left(D_j\right)P\left(\mathbf{k}/D_j\right)}, \qquad i = 0, 1, \ldots, N. \qquad \text{(Л19.6)}$$

Вероятности, входящие в обобщенную формулу Байеса, определяются из очевидных соображений:

$$P\left(D_i\right) = N_i/N\,;\ \sum_{i=0}^{N} P\left(D_i\right) = 1\,;\ P\left(\mathbf{k}/D_i\right) = N_{ik}/N\,;\ P\left(\mathbf{k}\right) = N_k/N\,. \qquad \text{(Л19.7)}$$

Здесь $N$ — общее число продиагностированных объектов, из которых $N_i$ — число объектов с диагнозом $D_i$; $N_{ik}$ — количество объектов с диагнозом $D_i$, имеющих комплексный признак $\mathbf{k}$; $N_k$ — число объектов во всех состояниях с признаком $\mathbf{k}$.

Если компоненты комплексного признака статистически независимы, то согласно формулам теории вероятности

$$P(\mathbf{k}/D_i) = \prod_{p=1}^{m} P(k_p/D_i). \tag{Л19.8}$$

В ином случае

$$P(\mathbf{k}/D_i) = P(k_1, k_2, ..., k_m/D_i) = P(k_1/D_i)P(k_2/k_1D_i)\cdot ...\cdot P(k_m/k_1, k_2, ..., k_{m-1}D_i).$$

Обобщенная формула Байеса в случае независимости компонентов комплексного признака записывается в виде

$$P(D_i/\mathbf{k}) = \frac{P(D_i)\prod_{p=1}^{m} P(k_p/D_i)}{\sum_{j=0}^{N} P(D_j)\prod_{p=1}^{m} P(k_p/D_j)}. \tag{Л19.9}$$

**Задача Л19.1.** Для диагностики подшипника главного циркуляционного насоса ЯЭУ используются два диагностических признака: $k_1$ — повышение температуры подшипника на 15 °С и более; $k_2$ — повышение уровня вибрации на 6 дБ и более. Появление этих признаков может быть связано с неисправностями в системе смазки подшипника — состояние $D_1$, либо с износом деталей подшипника — состояние $D_2$. Известно, что при нормальном состоянии подшипника (состояние $D_0$) признак $k_1$ не наблюдается, а признак $k_2$ наблюдается в 1 % случаев. На этапе предварительных исследований также установлено, что 90 % подшипников вырабатывают ресурс в нормальном состоянии, 2 % подшипников — в состоянии $D_1$, а 8 % — в состоянии $D_2$. Кроме того, известно, что признак $k_1$ встречается в состоянии $D_1$ в 60 % случаях, а в состоянии $D_2$ в 20 %. Признак $k_2$ в состоянии $D_1$ встречается в 30 %, а в состоянии $D_2$ — в 80 % случаях. Необходимо установить вероятности состояний подшипника при различных сочетаниях наблюдаемых признаков и их отсутствии, полагая признаки независимыми, то есть один из них может наблюдаться независимо от другого.

Исходные данные сведем в табл. Л19.1. Так как признак $k_j$ $(j = 1, 2)$ либо наблюдается, либо не наблюдается — событие $\overline{k}_j$, то при диагностике подшипника может быть установлен один из четырех комплексных признаков, компонентами которых являются $k_j$ и $\overline{k}_j$, $(j=1,2)$: $\mathbf{k_1}=(k_1;k_2)$; $\mathbf{k_2}=(\overline{k_1};k_2)$; $\mathbf{k_3}=(k_1;\overline{k_2})$; $\mathbf{k_4}=(\overline{k_1};\overline{k_2})$.

Таблица Л19.1. **Вероятности признаков и априорные вероятности появлений состояний подшипника**

| $D_i$ | $P(k_1/D_i)$ | $P(k_2/D_i)$ | $P(D_i)$ |
|---|---|---|---|
| $D_0$ | 0,0 | 0,01 | 0,90 |
| $D_1$ | 0,60 | 0,30 | 0,02 |
| $D_2$ | 0,20 | 0,80 | 0,08 |

Поскольку наблюдение признака $k_j$ и его отсутствие $\overline{k_j}$ — взаимоисключающие события, условная вероятность отсутствия признака определяется равенством

$$P\left(\overline{k_j}/D_i\right)=1-P\left(k_j/D_i\right). \tag{Л19.10}$$

Применяя обобщенную формулу Байеса (Л19.6) с учетом соотношений (Л19.8) и (Л19.10), определяем апостериорные вероятности диагнозов при наличии указанных сочетаний признаков. Например,

$$P(D_0/\mathbf{k_1})=\frac{P(D_0)P(k_1/D_0)P(k_2/D_0)}{P(D_0)P(k_1/D_0)P(k_2/D_0)+P(D_1)P(k_1/D_1)P(k_2/D_1)+}\rightarrow$$

$$\rightarrow\frac{}{+P(D_2)P(k_1/D_2)P(k_2/D_2)}=$$

$$=\frac{0{,}9\cdot 0{,}0\cdot 0{,}01}{0{,}9\cdot 0{,}0\cdot 0{,}01+0{,}02\cdot 0{,}6\cdot 0{,}3+0{,}08\cdot 0{,}2\cdot 0{,}8}=0;$$

$$P(D_1/\mathbf{k_1})=\frac{P(D_1)P(k_1/D_1)P(k_2/D_1)}{P(D_0)P(k_1/D_0)P(k_2/D_0)+P(D_1)P(k_1/D_1)P(k_2/D_1)+}\rightarrow$$

$$\rightarrow\frac{}{+P(D_2)P(k_1/D_2)P(k_2/D_2)}=$$

$$=\frac{0{,}02\cdot 0{,}6\cdot 0{,}3}{0{,}9\cdot 0{,}0\cdot 0{,}01+0{,}02\cdot 0{,}6\cdot 0{,}3+0{,}08\cdot 0{,}2\cdot 0{,}8}=0{,}22;$$

$$P\left(D_0/\mathbf{k}_2\right)=\frac{P\left(D_0\right)P\left(\overline{k}_1/D_0\right)P\left(k_2/D_0\right)}{P\left(D_0\right)P\left(\overline{k}_1/D_0\right)P\left(k_2/D_0\right)+P\left(D_1\right)P\left(\overline{k}_1/D_1\right)P\left(k_2/D_1\right)+}\rightarrow$$

$$\rightarrow\frac{}{+P\left(D_2\right)P\left(\overline{k}_1/D_2\right)P\left(k_2/D_2\right)}=$$

$$=\frac{0,9\cdot\left(1-0,0\right)\cdot 0,01}{0,9\cdot\left(1-0,0\right)\cdot 0,01+0,02\cdot\left(1-0,6\right)\cdot 0,3+0,08\cdot\left(1-0,2\right)\cdot 0,8}=0,14;$$

и т. д. Результаты расчетов сведены в табл. Л19.2.

Таблица Л19.2. **Апостериорные вероятности состояний подшипника**

| $D_i$ | $P\left(D_i/\mathbf{k}_1\right)$ | $P\left(D_i/\mathbf{k}_2\right)$ | $P\left(D_i/\mathbf{k}_3\right)$ | $P\left(D_i/\mathbf{k}_4\right)$ |
|---|---|---|---|---|
| $D_0$ | 0,00 | 0,14 | 0,00 | 0,98 |
| $D_1$ | 0,22 | 0,04 | 0,72 | 0,01 |
| $D_2$ | 0,78 | 0,82 | 0,28 | 0,01 |

Из данных, приведенных в табл. Л19.2, следует, что даже в отсутствие признаков $k_1$ и $k_2$, то есть при наличии комплексного признака $\mathbf{k_4}=\left(\overline{k}_1;\overline{k}_2\right)$, вероятности состояний $D_1$ и $D_2$ — $P\left(D_1/\mathbf{k_4}\right)$ и $P\left(D_2/\mathbf{k_4}\right)$, отличны от нуля. Это означает, что неисправности в системе смазки и износ деталей подшипника могут и не вызвать повышение температуры и уровня вибрации подшипника. Другими словами, эти признаки не являются достаточными для постановки диагнозов $D_1$ и $D_2$. Хотя вероятности таких событий малы.

При наличии обоих признаков достоверно можно утверждать, что подшипник неисправен. Наиболее вероятным состоянием является износ деталей подшипника — состояние $D_2$. В отсутствие признаков наиболее вероятно исправное состояние (вероятность его равна 0,98). При наличии признака $k_2$ и отсутствии $k_1$ вероятность износа — состояния $D_2$ больше. При наличии признака $k_1$ и отсутствии $k_2$ больше вероятность неисправности в системе смазки — состояние $D_1$. Если задать уровень распознавания $P^*=0,9$, то следует констатировать, что для достоверного определения типа неисправности подшипника недостаточно данных, и необходимы дополнительные исследования.

## Метод последовательного анализа

Метод последовательного анализа, или метод Вальда [2], используется в основном при *дихотомии*, то есть тогда, когда по результатам диагностики необходимо принять решение о нахождении объекта в одном из двух возможных состояний: $D_0$ — исправное состояние и $D_1$ — неисправное состояние. В отличие от метода Байеса при последовательном анализе проводится последовательное обследование диагностируемого объекта по компонентам сложного признака $\mathbf{k}$. После этого уточняются вероятности состояний, и обследование прекращается, когда достигается заданная вероятность правильной постановки диагноза.

Для пояснения сущности метода Вальда рассмотрим отношение апостериорных вероятностей диагнозов $D_0$ и $D_1$ при наличии у объекта сложного признака $\mathbf{k}$, состоящего из независимых компонент. Прежде всего, отметим, что в случае дихотомии по методу Байеса согласно (Л19.1) ставится диагноз $D_0$, если

$$\frac{P(D_0)}{P(D_1)} \cdot \frac{\mathrm{P}(\mathbf{k}/D_0)}{P(\mathbf{k}/D_1)} > 1. \qquad \text{(Л19.11)}$$

В ином случае — диагноз $D_1$. Отношение условных вероятностей обнаружить у диагностируемого объекта сложный признак $\mathbf{k}$, если объект находится в состояниях $D_0$ и $D_1$:

$$\Lambda(\mathbf{k}) = \Lambda(k_1,\ \ldots,\ k_m) = \frac{P(\mathbf{k}/D_0)}{P(\mathbf{k}/D_1)}$$

называется *отношением правдоподобия*. Правило (Л19.11) можно выразить через отношение правдоподобия. При

$$\Lambda(\mathbf{k}) > \frac{P(D_1)}{P(D_0)} \qquad \text{(Л19.12)}$$

ставится диагноз $D_0$, если выполняется обратное неравенство — $D_1$.

По методу Вальда проводится последовательное обследование объекта по компонентам сложного признака. На первом этапе определяется отношение правдоподобия по первому диагностическому признаку $k_1$ и выбираются верхние и нижние границы принятия решения $A$ и $B$.

Если

$$\Lambda\left(k_1\right)=\frac{P\left(k_1/D_0\right)}{P\left(k_1/D_1\right)}>A\,, \qquad (Л19.13)$$

делается заключение о состоянии $D_0$.

Если

$$\Lambda\left(k_1\right)=\frac{P\left(k_1/D_0\right)}{P\left(k_1/D_1\right)}<B \qquad (Л19.14)$$

ставится диагноз $D_1$.

Если же

$$B<\Lambda\left(k_1\right)<A, \qquad (Л19.15)$$

информации недостаточно для определения состояния объекта, и обследование проводится по следующей компоненте $k_2$. Процедура продолжается до тех пор, пока на каком-то этапе не будет выполнено одно из неравенств вида (Л19.13) или (Л19.14). Если в результате обследования все компоненты сложного признака исчерпаны, но ни одно из неравенств (Л19.13) и (Л19.14) не выполняется, отказываются от постановки диагноза — диагностических признаков недостаточно для определения состояния объекта.

Для удобства вычислений обычно рассматривают не отношение правдоподобия, а его монотонную функцию–логарифм. Тогда условие постановки диагнозов $D_0$ и $D_1$ после *l*-го обследования принимает соответственно вид

$$\ln\left\{\Lambda\left(k_1,k_2,...,k_l\right)\right\}=\sum_{i=1}^{l}\ln\left\{P\left(k_i/D_0\right)\big/P\left(k_i/D_1\right)\right\}>\ln A\,;$$

$$\ln\left\{\Lambda\left(k_1,k_2,...,k_l\right)\right\}=\sum_{i=1}^{l}\ln\left\{P\left(k_i/D_0\right)\big/P\left(k_i/D_1\right)\right\}<\ln B\,. \qquad (Л19.16)$$

Заметим, что если обследование проводится по диагностическим параметрам, являющимся непрерывными случайными величинами, то в вышеприведенных соотношениях надо условные вероятности распределения признаков заменить на условные плотности вероятности распределения параметров.

При постановке диагноза возможны ошибочные решения. Если в результате диагностики принято решение, что объект находится в неисправном состоянии $D_1$, когда, на самом деле, он исправен (состояние $D_0$), считается, что допущена ошибка 1-го рода. Ее называют еще ложной тревогой. Вероятность ошибки 1-го рода обозначим через $P_F$.

Если принимается решение, что объект исправен, то есть находится в состоянии $D_0$, когда его действительное состояние неисправное $D_1$, допускается ошибка 2-го рода. Другое ее название — пропуск сигнала. Вероятность ошибки 2-го рода будем обозначать как $P_M$.

Ошибочные решения о состоянии диагностируемого объекта, как правило, приводят к потерям. Если допускается ошибка 2-го рода, и исправный объект выводится из эксплуатации, то потери могут быть связаны с затратами на ремонтные работы, в которых нет необходимости, или с недополученной выручкой от реализации продукции, производство которой приостановлено. Эти потери, выраженные в денежных единицах, обозначим через $\Pi_{01}$. В случае ошибки 2-го рода потери могут быть обусловлены затратами на ликвидацию аварии и ее последствий, если допущенный к дальнейшей эксплуатации неисправный технический объект выходит из строя, а также недополученными средствами от продажи продукции из-за остановки производства. Потери, связанные с ошибкой 2-го рода, обозначим $\Pi_{10}$. Если можно оценить выгоду при правильной постановке диагноза, то вводят выигрыши $\Pi_{00} < 0$ и $\Pi_{11} < 0$ — отрицательные потери. Если определить выигрыши невозможно, их полагают равными нулю. Названия, обозначения и стоимости (потери) от ошибочных решений при определении состояния диагностируемых объектов, указаны в табл. Л19.3.

Таблица Л19.3. **Ошибки и их стоимости при определении состояния технических объектов**

| Поставленный диагноз | Истинное состояние | Название ошибки | Вероятность ошибки | Обозначения потерь при принятии решений |
|---|---|---|---|---|
| $D_0$ | $D_0$ | — | — | $\Pi_{00}$ |
| $D_1$ | $D_1$ | — | — | $\Pi_{11}$ |
| $D_1$ | $D_0$ | 1-го рода (ложная тревога) | $P_F$ | $\Pi_{01}$ |
| $D_0$ | $D_1$ | 2-го рода (пропуск сигнала) | $P_M$ | $\Pi_{10}$ |

Если принимается решение о диагнозе $D_0$, то вероятность правильного решения равна $1 - P_M$. Вероятность отнесения объекта к состоянию $D_1$ в этом случае будет равна $P_F$. В силу (Л19.14) вероятность диагноза $D_0$ по крайней мере в $A$ раз превышает вероятность диагноза $D_1$, следовательно,

$$\frac{1-P_M}{P_F} \geq A. \tag{Л19.17}$$

Аналогично можно получить следующую оценку

$$\frac{P_M}{1-P_F} \leq B. \tag{Л19.18}$$

Соотношения (Л19.17) и (Л19.18), рассматриваемые как строгие равенства, позволяют определить верхнюю и нижнюю границы принятия решения $A$ и $B$ в (Л19.13) – (Л19.15) при заданных вероятностях ошибок первого и второго рода $P_F$ и $P_M$ [2].