

**Figure 5.9** The confusion matrix for the cancer treatment problem, in which the rows correspond to the true class and the columns correspond to the assignment of class made by our decision criterion. The elements of the matrix show the numbers of true negatives, false positives, false negatives, and true positives.

	normal	cancer
normal	$N_{TN}$	$N_{FP}$
cancer	$N_{FN}$	$N_{TP}$

### Section 11.2.3

## Chapter 7

the resulting posterior probabilities so they sum to one. The particular conditional independence assumption (5.26) is an example of a *naive Bayes model*. Note that the joint marginal distribution  $p(\mathbf{x}_I, \mathbf{x}_B)$  will typically not factorize under this model. We will see in later chapters how to construct models for combining data that do not require the conditional independence assumption (5.26). A further advantage of using models that output probabilities rather than decisions is that they can easily be made differentiable with respect to any adjustable parameters (such as the weight coefficients in the polynomial regression example), which allows them to be composed and trained jointly using gradient-based optimization methods.

### 5.2.5 Classifier accuracy

The simplest measure of performance for a classifier is the fraction of test set points that are correctly classified. However, we have seen that different types of error can have different consequences, as expressed through the loss matrix, and often we therefore do not simply wish to minimize the number of misclassifications. By changing the location of the decision boundary, we can make trade-offs between different kinds of error, for example with the goal of minimizing an expected loss. Because this is such an important concept, we will introduce some definitions and terminology so that the performance of a classifier can be better characterized.

### Section 2.1.1

We will consider again our cancer screening example. For each person tested, there is a ‘true label’ of whether they have cancer or not, and there is also the prediction made by the classifier. If, for a particular person, the classifier predicts cancer and this is in fact the true label, then the prediction is called a *true positive*. However, if the person does not have cancer it is a *false positive*. Likewise, if the classifier predicts that a person does not have cancer and this is correct, then the prediction is called a *true negative*, otherwise it is a *false negative*. The false positives are also known as *type 1 errors* whereas the false negatives are called *type 2 errors*. If  $N$  is the total number of people taking the test, then  $N_{TP}$  is the number of true positives,  $N_{FP}$  is the number of false positives,  $N_{TN}$  is the number of true negatives, and  $N_{FN}$  is the number of false negatives, where

$$N = N_{TP} + N_{FP} + N_{TN} + N_{FN}. \quad (5.28)$$

This can be represented as a *confusion matrix* as shown in Figure 5.9. Accuracy, measured by the fraction of correct classifications, is then given by

$$\text{Accuracy} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FP} + N_{TN} + N_{FN}}. \quad (5.29)$$