4.5 Linear Least Squares: Example and remark

Example: *straight line*

The model function for a straight line is $f(x) = ax + b$, which depends linearly on the two parameters a and b. Consider the eleven given data points shown in Fig. 4.3(a) that are to be fitted by a straight line. Here, the "model vector" can be written as

$$
\begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_{11} \end{pmatrix} = \overbrace{\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_{11} & 1 \end{pmatrix}}^{\vec{M}} \begin{pmatrix} a \\ b \end{pmatrix}, \tag{4.22}
$$

i.e., M is an 11×2 matrix with a column of constant values of 1 corresponding to parameter b. **Remark:** *robust*
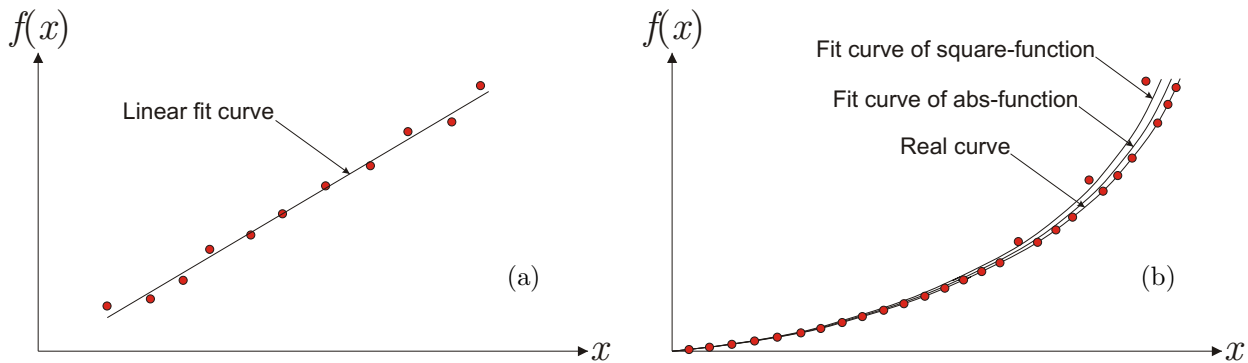

Figure 4.3: (a) Fitting a straight line to given data points. (b) Comparison of fitting results for the usual least squares minimization ("square function") and robust fitting using the absolute value of the differences ("abs function").

fitting

Obviously, since they minimize that function, the fitting parameter values obtained depend on the objective function used, which will influence the fitting result. This is of importance in the following situation: When fitting experimental data it may happen that a few data points lie noticeably far away from the expected model curve, which will lead to a large contribution of these points in the χ^2 function. As a consequence, the fitted curve will be drawn close to these points in order to minimize the gap. However, from the point of view of an experimentalist, it is not desirable that a few points that can be considered an exception have a large influence on the result.

To reduce the influence of data points with large deviations from the model curve, a linear objective function can be used instead of the χ^2, which is called *robust fitting*:

$$
\chi_{\text{abs}}(\vec{a}) = \sum_{i=1}^{n} |y_i - f(x_i, \vec{a})|. \tag{4.23}
$$

The main difference between χ_{abs} and χ^2 (Eq. 4.15) is that χ^2 can be differentiated at all values while χ_{abs} does not have derivative everywhere. To illustrate the influence of using χ_{abs} instead of χ^2, Fig. 4.3(b) shows a constructed example of fitting a parabola where a few data points are shifted upwards. Obviously, the χ^2 result deviates stronger from the true parabola.