10 Validation

In Chapter 8 we introduced the mathematical model $\underline{y} = Ax + \underline{e}$, with $m \geq n$, where m is the dimension of vector y, and n the dimension of vector x. The system $y = Ax$ is generally not consistent, hence, $y \neq Ax$. The least-squares estimate (8.4), and the minimum variance estimator (8.6) were introduced, Sections 8.3.2 and 8.3.4. Once an estimate $\hat{x}$ for the unknown parameters is available, one can compute an estimate for the observations: $\hat{y} = A\hat{x}$, and this system is *consistent*. One could regard this parameter estimation as 'changing' (or adjusting) the observed values from y to $\hat{y}$ in order to turn a non-consistent system into a consistent one. As outlined with the least-squares criterion (8.5), one keeps vector $y - \hat{y}$ as short as possible, the observed values should be 'changed only as little as possible'.

In this chapter we introduce the least-squares residuals, and show how they can be used in an overall *consistency check*, to answer the question whether the collected measurements y and the assumed model Ax can be deemed to be mutually consistent. Next we present a worked example of line-fitting (regression). Eventually we briefly introduce the more advanced (optional) subject of observation testing, and present the final results of our data processing and analysis.

10.1. Least-squares residuals

Once an estimator $\underline{\hat{y}}$ is available for the vector of observables, the least-squares residuals follow as

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}} \tag{10.1}$$

The least-squares residuals measure the difference between the observations (as measured) y, and the estimated, or adapted ones $\hat{y}$ (see Section 8.3.2, $\underline{\hat{y}} = A\underline{\hat{x}}$). The least-squares residuals provide an estimate for the (unknown) measurement error e (that is why also the residuals are denoted with a hat-symbol). They carry important diagnostic information about the parameter estimation process. When the residuals are small, the situation is looking good. One does not need to 'change' the observed values by much, in order to make them fit into the model Ax. However, when they are large, this might be a reason for reconsideration. It could be that there are large outliers or faults present in the measurements (e.g. entering an observed height difference as 0.413 m, instead of 0.143 m), or that the assumed model is not appropriate for the case at hand (in a dynamic system, with a moving object, we may assume that the object is moving with constant velocity, but this may turn out not to be the case). Small residuals tend to be OK — large ones are not. But what is small, and what is big? Fortunately, we can devise an objective criterion to judge on their size.

10.1.1. Overall model test - consistency check

When the data are normally distributed, $\underline{y} \sim N(Ax, Q_{yy})$, under the working model (the null-hypothesis), then also the residuals will be normally distributed (with zero mean). It can be shown — though the proof is omitted here — that $\underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}}$ has — under the working model (8.2) — a central Chi-squared distribution with $m-n$ degrees of freedom, hence $\chi^2(m-n, 0)$ (from the m-vector of observations, n unknown parameters in x are estimated, and hence only $(m-n)$ degrees of freedom are 'left' for the residuals). The Chi-squared distribution was introduced in Section 9.5, and shown in Figure 9.8. The *squared norm of the residuals-vector* $\underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}}$ is an *overall* measure of consistency[1] [2]. It provides an objective criterion on judging the amount by which we needed to 'change' the observed values, to fit the assumed or supposed model (Ax).

$$\underline{T} = \underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}} \sim \chi^2(m-n, 0) \tag{10.2}$$

Now one can set a level of significance (probability) α, see Chapter 26 in [2], e.g. 5%, and not accept the residual vector $\hat{e}$, once its squared norm is located in the upper 5% of the distribution (right tail); occurrence of such a value is deemed to be too unlikely to be true under the working model. When there are large outliers or faults present in the observations, or when the assumed model is not appropriate, the residuals tend to be larger, leading to a *larger* value for the test-statistic T, and hence we set a right-sided critical region.

In practice, one computes $T = \hat{e}^T Q_{yy}^{-1} \hat{e}$, and retrieves threshold $k' = \chi^2_\alpha(m-n, 0)$ from the table and concludes (decides):

- model and data are consistent when $T < k'$
- model and data are *not* consistent when $T > k'$

For example with m=5 and n=2, $m-n$=3, and with a 5% level of significance, the threshold (or critical) value k' for the above squared norm of the least-squares residuals, T, is 7.8147, see the table in Appendix D ($\chi^2_\alpha(3, 0) = 7.8147$).

10.1.2. Simplification

In the simple case that the observables' variance matrix is a diagonal matrix $Q_{yy} = \text{diag}(\sigma^2_{y_1}, \sigma^2_{y_2}, \ldots, \sigma^2_{y_m})$ (all observables are uncorrelated), the above overall model test statistic $\underline{T}$ can be given a simple and straightforward interpretation. Namely

$$\underline{T} = \underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}} = \sum_{i=1}^{m} \frac{\underline{\hat{e}}_i^2}{\sigma^2_{y_i}} = \sum_{i=1}^{m} \left(\frac{\underline{\hat{e}}_i}{\sigma_{y_i}} \right)^2 \tag{10.3}$$

and it compares — per observation — the residual $\hat{e}_i$ with the standard deviation σ_{y_i} of, i.e. the expected uncertainty in, the observable $\underline{y}_i$.

The overall model test statistic equals the sum of all those m squared ratios. The overall model test, as the name says, aims to *detect*, in general sense, any inconsistency between observed data and the proposed or assumed model.

[1] actually a measure of in-consistency

[2] formally $\underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}}$ is the square of the *weighted* norm (weighted because of Q_{yy}^{-1}) of the least-squares residuals; in statistics unweighted $\underline{\hat{e}}^T \underline{\hat{e}}$ is referred to as sum of squared residuals (SSR)

10.1.3. Discussion

As an example we consider the case in which the same unknown distance is measured twice, cf. Section 8.3.1. If one of these measurements is biased by a large amount (for instance 10 m), and the other is not, then the overall model test will likely detect that the two measurements are not consistent with the model, namely, according to the model (measuring the same distance twice) the numerical values of two measurements should be the same, or close together. Intuitively, as the measurements are not close together, this gives rise to suspicion (there might be something wrong with the measurements). Anomalies in the measurements which cause the data to be (still) consistent with the assumed model can*not* be detected by this test. In our example, if both observations are biased in the same way, let us say by 10 m, then the two observations are still consistent (with each other in the model; they are both in error). Intuitively, as their values are the same or close together, this does not raise any suspicion. Being able to detect all relevant anomaly scenarios is part of designing a good measurement set-up.

10.1.4. Example: repeated measurements [*]

Suppose m measurements of the same unknown quantity are made. Then the general model of observation equations $E(\underline{y}) = Ax;\ D(\underline{y}) = Q_{yy}$ (8.2) reads:

$$E\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_m \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} x \ ; \ D\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_m \end{pmatrix} = \sigma^2 I_m$$

where we assumed that all observables have equal precision (all variances equal to σ^2), and there is no correlation. There are m observations, and there is $n = 1$ unknown parameter.

In this case, the minimum variance estimate $\hat{x}$ for the unknown parameter x, equals just the mean of the observations, similar to (6.12).

In this simple example $\hat{y}_i = \hat{x}$, and hence, the least-squares residuals $\hat{e}_i = y_i - \hat{x}$. The squared norm of the residuals vector becomes

$$T = \hat{e}^T Q_{yy}^{-1} \hat{e} = \frac{1}{\sigma^2} \sum_{i=1}^{m} \hat{e}_i^2 = \frac{1}{\sigma^2} \sum_{i=1}^{m} (y_i - \hat{x})^2$$

which shows that, in this case, this overal model test-statistic is closely related to the sample variance (6.14), which you determine based on measurements taken. We have

$$\frac{T}{m-1} = \frac{1}{\sigma^2} \frac{1}{m-1} \sum_{i=1}^{m} (y_i - \hat{x})^2 = \frac{\hat{\sigma}^2}{\sigma^2}$$

where we use m instead of N as in (6.14). The overall model test statistic equals, in this case, the ratio of the sample variance and the formal, a-priori variance.

10.2. Example

In this example we consider the classical problem of *line fitting*. This problem is frequently encountered in science and engineering. One can think of measuring the extension of a certain object (e.g. of steel) due to temperature, for which one assumes a linear behaviour, so the length of the object is measured, at different temperatures, and next one would like to determine the length of the object at some reference temperature and the coefficient of extension, i.e. by how much it extends for every increase of one degree in temperature. In literature, this subject is often referred to as regression analysis, see e.g. Chapter 6 on orthogonality and least-squares in [27], in particular Section 6 with applications to linear models. The subject can be seen much broader however, namely as *curve fitting*. The principle is not restricted to just straight lines, one can also use parabolas and higher degree polynomials for instance.

In this example we consider a vehicle which is driving along a straight line, and we are interested in the position along the road. Therefore, a laser tracker is used, and this device measures/reports the position of the vehicle every second. For convenience, the laser-tracker is at the origin of this one-dimensional coordinate system.

Over a period of four seconds, we take measurements: $y = (y_1, y_2, y_3, y_4)^T$, hence the vector of observations has dimension $m = 4$. Correspondingly at times t_1, t_2, t_3 and t_4, the unknown positions are $x(t_1)$, $x(t_2)$, $x(t_3)$ and $x(t_4)$. The measurements equal the unknown positions, apart from measurement errors, i.e. $y_i = x(t_i) + e_i$ for $i = 1, \dots, 4$.

$$E\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \end{pmatrix} = \begin{pmatrix} x(t_1) \\ x(t_2) \\ x(t_3) \\ x(t_4) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x(t_1) \\ x(t_2) \\ x(t_3) \\ x(t_4) \end{pmatrix}$$

In case we determine these $n = 4$ four unknown positions from the $m = 4$ four observed positions, estimation is pretty trivial, namely $\hat{x}(t_i) = y_i$ for $i = 1, \dots, 4$.

We have reasons however, to assume that the vehicle is driving at constant speed, and thereby we can model the unknown motion of the vehicle, by just an unknown initial position $x(t_0)$, and its velocity $\dot{x}$.

$$\begin{pmatrix} x(t_1) \\ x(t_2) \\ x(t_3) \\ x(t_4) \end{pmatrix} = \begin{pmatrix} 1 & (t_1 - t_0) \\ 1 & (t_2 - t_0) \\ 1 & (t_3 - t_0) \\ 1 & (t_4 - t_0) \end{pmatrix} \begin{pmatrix} x(t_0) \\ \dot{x} \end{pmatrix}$$

which we can substitute in the above system of observation equations, and hence

$$E\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & (t_1 - t_0) \\ 1 & (t_2 - t_0) \\ 1 & (t_3 - t_0) \\ 1 & (t_4 - t_0) \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x(t_0) \\ \dot{x} \end{pmatrix}}_{x}$$

Now there are still $m = 4$ observations, but just $n = 2$ unknown parameters in vector x. The distance observables by the laser tracker are all uncorrelated, and all have the same variance σ^2. Hence $Q_{yy} = \sigma^2 I_4$. Standard deviation σ can be taken here as $\sigma = 1$.

In the sequel we develop the example into a real numerical example. The observation times are $t_1 = 1$, $t_2 = 2$, $t_3 = 3$ and $t_4 = 4$ seconds, and $t_0 = 0$ (and timing is assumed here to be perfect — no errors; all coefficients of matrix A are known — when this assumption cannot be made, refer to Appendix B.6).

As can be seen in Figure 10.1, we try to fit a straight line through the observed data points. Therefore we estimate the offset/intercept of the line $x(t_0)$, and its slope $\dot{x}$; in terms of regression, they are the (unknown) regression coefficients. In this example time t is the explanatory or independent variable (also called the regressor); the (observed) position depends on time t. And, the observation y is the dependent (or response) variable. Least-squares estimation will yield a best fit of the line with the observed data points, through minimizing (8.5).

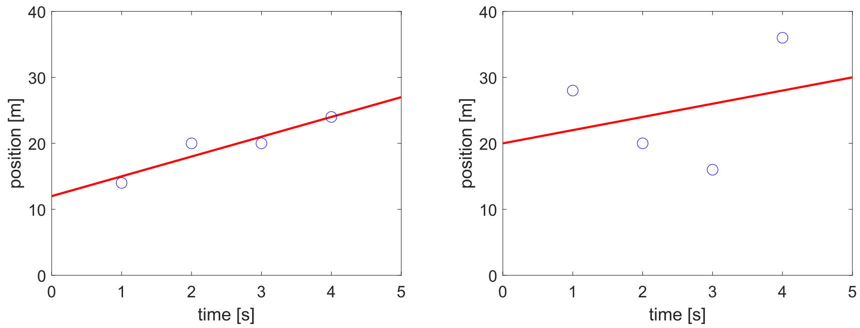

Figure 10.1: Line fitting on day 1 (left) and day 2 (right). Shown are the observed data points as blue circles, and the straight line in red, fitted by means of least-squares.

The experiment was repeated the next day, and in the sequel we consider — simultaneously — the outcomes of *both* experiments. The measurements on day 1 were $y = (14, 20, 20, 24)^T$, and on day 2 $y = (28, 20, 16, 36)^T$. In Figure 10.1, the observed distances (in meters) are shown, together with the fitted line, based on the least-squares estimates. Verify yourself that, with (8.6), one obtains

$$\hat{x} = \begin{pmatrix} 12 \\ 3 \end{pmatrix} \text{for day 1, and } \hat{x} = \begin{pmatrix} 20 \\ 2 \end{pmatrix} \text{for day 2}$$

with units in meters and meters per second respectively.

With $\hat{x}(t_0) = 12$ and $\hat{\dot{x}} = 3$ on day 1, one can determine $\hat{y}$ through $\hat{y} = A\hat{x}$, and eventually obtain $\hat{e} = y - \hat{y}$. And do this for day 2 as well.

$$\hat{e} = \begin{pmatrix} -1 \\ 2 \\ -1 \\ 0 \end{pmatrix} \text{ for day 1, and } \hat{e} = \begin{pmatrix} 6 \\ -4 \\ -10 \\ 8 \end{pmatrix} \text{ for day 2}$$

From these numbers, and also Figure 10.1, one can already conclude that for day 1 (on the left) we have a pretty good fit, while on day 2 (on the right) the fit is pretty poor, likely a second degree polynomial would do here much better (quadratic polynomial). It indicates that the motion of the vehicle on day 2 has not really been at constant speed. For an objective criterion in judging a good and poor fit, we use the squared norm of the residual vector (10.2).

$$T = \hat{e}^T Q_{yy}^{-1} \hat{e} = 6 \text{ for day 1, and } T = \hat{e}^T Q_{yy}^{-1} \hat{e} = 216 \text{ for day 2}$$

With the table in Appendix D, we find that with $m - n = 2$ and $\alpha = 0.01$ the threshold value equals $\chi^2_\alpha = 9.2103$, and hence the line fit of day 1 is not rejected ($6 < 9.2103$), but the line fit of day 2 is rejected ($216 > 9.2103$)!

At this stage we recall that least-squares estimation is driven, see (8.5), by minimization of the squared norm of the vector $y - Ax$, see Figure 10.2. Least-squares minimizes the objective function $\min_x \|y - Ax\|^2$. Vector $y - Ax$ contains the differences between the actual observations y and the modeled observations by Ax. One should choose values for the elements in vector x, such that the squared norm of th vector $y - Ax$ is at minimum. Figure 10.2 shows at left, for the observations of day 1, along the vertical axis, the squared norm of the vector $y - Ax$, which is just a single number, as a function of $x(t_0)$ and $\dot{x}$ in the horizontal plane, with offset parameter $x(t_0)$ along the axis in front, and slope parameter $\dot{x}$ along the axis to the back.

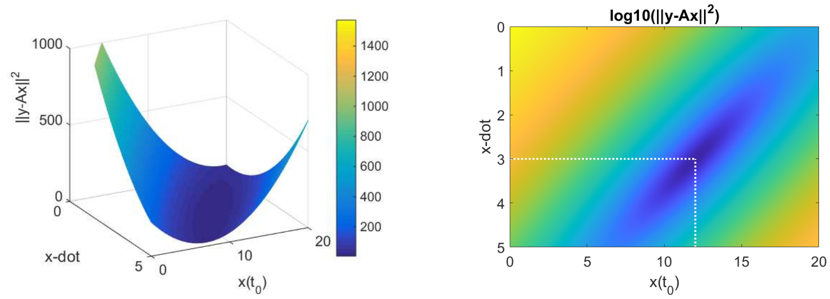

Figure 10.2: The sum of squared residuals as a function of $x(t_0)$ and $\dot{x}$. At right, actually the $\log 10$-value is shown. The minimum, which is equal to 6, is attained for $x(t_0)$=12 and $\dot{x}$=3, indicated by the dotted lines, hence $\min_x \|y - Ax\|^2$ = 6 (for day 1).

We have to keep in mind that the true motion of the vehicle is unknown. It is *our assumption* that we can mathematically describe the motion through a constant velocity model. For day 1 there is no indication that this is not working. But for day 2, a constant velocity model, most likely, is not providing a proper description of the actual motion. Maybe a quadratic model (constant acceleration) gives a better/acceptable fit.

$$E\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \end{pmatrix} = \begin{pmatrix} 1 & (t_1 - t_0) & \frac{1}{2}(t_1 - t_0)^2 \\ 1 & (t_2 - t_0) & \frac{1}{2}(t_2 - t_0)^2 \\ 1 & (t_3 - t_0) & \frac{1}{2}(t_3 - t_0)^2 \\ 1 & (t_4 - t_0) & \frac{1}{2}(t_4 - t_0)^2 \end{pmatrix} \begin{pmatrix} x(t_0) \\ \dot{x}(t_0) \\ \ddot{x} \end{pmatrix}$$

There are now $n = 3$ unknown parameters, with still $m = 4$ observations.

Mind that if we go for a third degree polynomial (cubic polynomial), we will have $n = 4$ unknown parameters, and all information contained in the $m = 4$ observations is strictly needed to determine the unknown parameters, and 'nothing will be left' for the least-squares residuals, as $m - n = 0$. Without redundancy, the overall model test gets void. In that case it looks like the observations perfectly fit the assumed model, but you actually do not have any means to verify this.

10.3. Observation testing and outlook [*]

The squared norm of the residuals is an overall measure of consistency between observations and assumed mathematical model. The statistical test of $\hat{e}^T Q_{yy}^{-1} \hat{e}$ being smaller or larger than the critical value k' from the Chi-squared distribution is referred to as the overall model test, see Section 10.1.1. In literature you may also encounter it as the F-test, then $\underline{\hat{e}}^T Q_{yy}^{-1} \underline{\hat{e}}/(m - n) \sim F(m - n, \infty, 0)$ under the working model (null-hypothesis), where F represents the F-distribution.

The test can be derived from the principle of statistical hypothesis testing. More specific statistical hypothesis tests exist, for instance tests to identify outliers, blunders, faults and anomalies in single observations. A true coverage of this subject is beyond the scope of this book, but we will introduce — without any derivation, or proof of optimality — a simple test which aims to identify an *outlier* in a set of observations (then only *one* observation from this set is affected by the outlier). Typically this test is used multiple times, namely to test *each* of the m observations in vector $y = (y_1, \dots, y_m)^T$ separately. It is based again on the least-squares residuals $\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}$. Using the error propagation law (7.9), with (8.6), and $\underline{\hat{y}} = A\underline{\hat{x}}$, one can derive that

$$Q_{\hat{e}\hat{e}} = Q_{yy} - A(A^T Q_{yy}^{-1} A)^{-1} A^T = Q_{yy} - A Q_{\hat{x}\hat{x}} A^T$$

where $Q_{\hat{x}\hat{x}}$ was given by (8.7).

Under the working model the least-squares residuals have zero mean, hence $\underline{\hat{e}} \sim N(0, Q_{\hat{e}\hat{e}})$, and the vector $\underline{\hat{e}}$ is normally distributed as it is a linear function of $\underline{y}$, which is also taken to be normally distributed.

The least-squares residuals $\underline{\hat{e}}$ is a vector with m random variables $\underline{\hat{e}} = (\underline{\hat{e}}_1, \underline{\hat{e}}_2, \ldots, \underline{\hat{e}}_m)^T$. For each of the residuals we have $\underline{\hat{e}}_i \sim N(0, \sigma^2_{\hat{e}_i})$ with $i = 1, \ldots, m$. Usually, if (just) observation y_i contains a large error, the corresponding residual $\hat{e}_i$ will deviate (substantially) from zero. We use this to propose — valid only for the case when the observables have a diagonal variance matrix Q_{yy} — the w-test statistic as

$$\underline{w}_i = \frac{\underline{\hat{e}}_i}{\sigma_{\hat{e}_i}} \tag{10.4}$$

and check whether it deviates from zero.

Division of the residual by the standard deviation $\sigma_{\hat{e}_i}$ (you may use (7.9) again) causes the w-test statistic to be *standard normally distributed*: $\underline{w}_i \sim N(0, 1)$. In fact, it is the *normalized* residual. Setting a level of significance α, one can find the critical value. Mind that deviations both in positive and negative direction may occur. Hence, the level of significance α is split into $\frac{\alpha}{2}$ for the right tail, and $\frac{\alpha}{2}$ for the left tail. The critical value $\tilde{k} = N_{\frac{\alpha}{2}}(0, 1)$ follows from the table in Appendix C. The hypothesis 'no outlier present in observation y_i' is rejected if $|w_i| > \tilde{k}$, i.e. if $w_i < -\tilde{k}$ or $w_i > \tilde{k}$. The test is typically done for all observations $i = 1, \ldots, m$.

In practice the largest of the w-teststatistics (in absolute sense) indicates the most likely faulty observation. This observation is removed from the data, and estimation and validation is repeated with one less observation, until no more measurements are rejected, or until redundancy runs low. This procedure of 'data checking' can be carried out automatedly, without human intervention, and allows for a robust processing of the measurements. This checking provides a safeguard against producing, unknowingly, largely incorrect results.

There are many alternative approaches to robust estimation, delivering robust estimators, which are insensitive to outliers, but these are beyond the scope of this book.

10.4. Example — continued

Eventually we wrap up, and present the final results of our data processing and analysis. We have computed the estimates for the parameters of interest in the example of Section 10.2, namely the initial position $x(t_0)$ and the velocity $\dot{x}$, and we have verified the consistency of the model and the data.

In Figure 10.3 we consider the data of day 1, hence giving a good fit, see Figure 10.1 at left. In Figure 10.3 at left, we present again the four measurements, and, we indicate with them the measurement uncertainty. The standard deviation of each measurement was $\sigma = 1$. The graph presents the so-called error-bars for a confidence level of 95%. When the observables are normally distributed, one can verify with the table in Appendix C, that the error-bar extends by $r_\alpha \sigma$ to either side of the observed value, with r_α = 1.96.

The graph at right presents again the fitted line, based on the least-squares estimate $\hat{x}$. As outlined in the introduction of this chapter, estimates for the observations can be computed as $\hat{y} = A\hat{x}$. With this, you can check yourself that while the observation y_2 = 20, $\hat{y}_2$ = 18 (and the latter is located on the red line).

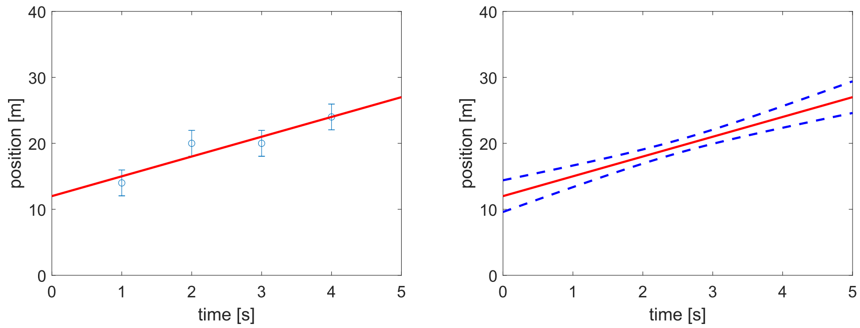

Figure 10.3: Line fitting on day 1: observed data points as small blue circles, with error-bars, also in blue, at the left, and the straight line fitted by means of least-squares with confidence interval at the right.

The corresponding variance matrix of $\underline{\hat{y}}$ can be found, using (7.9) on $\underline{\hat{y}} = A\underline{\hat{x}}$ as (7.7), as $Q_{\hat{y}\hat{y}} = AQ_{\hat{x}\hat{x}}A^T$, with $Q_{\hat{x}\hat{x}}$ from (8.7). So, for the estimated observation at t_2 holds that $\sigma_{\hat{y}_2} \approx 0.55$, whereas $\sigma_{y_2} = 1$ (hence, it is much smaller). Then the 95% confidence interval for the true y_2 is centered at $\hat{y}_2$, and extends to $r_\alpha \sigma_{\hat{y}_2} \approx 1.07$ on either side (and hence it is much smaller than with the individual observation, in the graph at left). Suppose the experiment is repeated, then 95% of the realizations of this (random) interval will actually contain the true position, see also Chapter 23 in [2]. In fact, the confidence interval can be presented for the position at any time t, see the dashed blue lines on either side of the fitted line, in the graph at right; it is the confidence interval for $y(t)$ based on $\hat{y}(t)$, for any time t.

10.5. Exercises and worked examples

This section presents several problems with worked answers on parameter estimation and validation.

Question 1 To determine the sea level height in the year 2000, and also its rate of change over time, five observations of the sea level are available, for instance from a tide gauge station. The observations are:

- y_1 = 25 mm, in the year 1998
- y_2 = 24 mm, in the year 1999
- y_3 = 27 mm, in the year 2000
- y_4 = 26 mm, in the year 2001
- y_5 = 28 mm, in the year 2002

Based on these five observations, compute the *least-squares* estimates for the sea level in the year 2000 and the rate of change. All quantities are referenced to the start of the year. The rate of change can be considered constant over the time span considered.

Answer 1 Similar to the example of Section 10.2, the model reads

$$E\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \underline{y}_3 \\ \underline{y}_4 \\ \underline{y}_5 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x(t_0) \\ \dot{x} \end{pmatrix}}_{x}$$

and basically we need to fit a straight line through five data points. In the above $y = Ax$ model, $x(t_0)$ is the *unknown* sea level height in the year 2000 (expressed in [mm]), and $\dot{x}$ is the (assumed constant) rate of change of the sea level (and also unknown, and expressed in [mm/year]). On the left hand side we have the five observations, of which the third, y_3, is the *observed* sea level height in the year 2000 (in [mm]). For example, the last observation, y_5, is related to the two unknown parameters as: y_5 equals (on average) the sum of the sea level in the year 2000 $x(t_0)$, plus twice the yearly change $\dot{x}$. These two coefficients, 1 and 2, show up, as the last row, in the A-matrix. There is no information given with regard to the precision of the observables (no matrix Q_{yy}), hence we use the basic least-squares equation (8.4) $\hat{x} = (A^T A)^{-1} A^T y$.

$$A^T A = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \quad (A^T A)^{-1} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{10} \end{pmatrix}$$

and $\hat{x} = (A^T A)^{-1} A^T y$ becomes

$$\hat{x} = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{10} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 25 \\ 24 \\ 27 \\ 26 \\ 28 \end{pmatrix} = \begin{pmatrix} 26 \\ \frac{8}{10} \end{pmatrix}$$

The least-squares estimate for the sea level height in the year 2000 is 26 mm, and the rate of change is 0.8 mm/year. Note that the least-squares estimate for the height in the year 2000 does *not* equal the observed height (y_3). The least-squares estimate is determined based on *all* available observations. In an era of climate change and sea-level rise, conscientious and responsible analysis and interpretation of sea-level height measurements over time is essential to design and maintenance of coastal defense infrastructure against flooding, see Figure 10.4.

Figure 10.4: The Oosterscheldekering, a series of dams and storm surge barriers to protect the province of Zeeland from flooding from the North-Sea. Photo by Rijkswaterstaat, 2007, taken from Beeldbank Rijkswaterstaat, under BY-NC license [28].

Question 2 With the model and observations of the previous question, determine whether the overall model test is passed or not, when the level of significance is set to 10%. The observables can be assumed to be all uncorrelated, and have a standard deviation of 1 mm (which is not really a realistic value in practice, but it is fine for this exercise).

Answer 2 The variance matrix of the observables reads $Q_{yy} = I_5$, and this does not change anything to the computed least-squares estimates (see Eq. (8.6)). The overall model test is

$T = \hat{e}^T Q_{yy}^{-1} \hat{e}$, hence we need the vector of least-squares residuals (10.1) $\hat{e} = y - \hat{y} = y - A\hat{x}$.

$$\begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \\ \hat{e}_4 \\ \hat{e}_5 \end{pmatrix} = \begin{pmatrix} 25 \\ 24 \\ 27 \\ 26 \\ 28 \end{pmatrix} - \begin{pmatrix} 24.4 \\ 25.2 \\ 26.0 \\ 26.8 \\ 27.6 \end{pmatrix} = \begin{pmatrix} 0.6 \\ -1.2 \\ 1.0 \\ -0.8 \\ 0.4 \end{pmatrix}$$

Then the value for the overall model test statistic becomes $T = \hat{e}^T Q_{yy}^{-1} \hat{e}$ = 3.6. The threshold, with $\alpha = 0.1$, is $k' = \chi^2_\alpha(m-n, 0)$, and $m = 5$ and $n = 2$, hence $m - n = 3$. With the table in Appendix D we obtain $k' = \chi^2_{0.1}(3,0)$ = 6.2514, and hence $T < k'$, and the overall model test is accepted. There is no reason to suspect that something is wrong; the assumed model and the made observations seem to be in agreement with each other, they seem to be consistent; the fit is good.

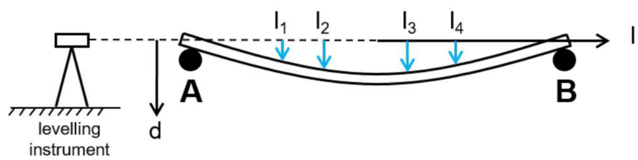

Figure 10.5: The deflection d of a bridge deck is observed at four positions l_1, l_2, l_3, and l_4.

i	l_i	$d(l_i)$
1	-2	1
2	-1	4
3	1	3
4	2	2

Table 10.1: Measurements of deflection $d(l_i)$ at four positions l_1, l_2, l_3, and l_4 at the bridge deck.

Question 3 A bridge deck is supported by pillars at points A and B. By its own weight, the deck in between will bend (deflect) as shown in Figure 10.5. A vertical cross-section along the bridge deck center-line, i.e. coordinate-axis l, is shown. The deflection d is modeled as a *quadratic* function of coordinate l: $d(l) = x_1 l^2 + x_2 l + x_3$, with unknown coefficients x_1, x_2, and x_3. Using leveling, the downward deflection d (downward is positive) has been observed at four positions along the deck. The measurements are listed in Table 10.1. Set up the model of observation equations $E(\underline{y}) = Ax$, for the four observations, for the goal of estimating the unknown coefficients x_1, x_2, and x_3.

Answer 3 We have four observation equations, according to the given quadratic function $d(l) = x_1 l^2 + x_2 l + x_3$. Hence,

$$E\begin{pmatrix} \underline{d}(l_1) \\ \underline{d}(l_2) \\ \underline{d}(l_3) \\ \underline{d}(l_4) \end{pmatrix} = \begin{pmatrix} l_1^2 & l_1 & 1 \\ l_2^2 & l_2 & 1 \\ l_3^2 & l_3 & 1 \\ l_4^2 & l_4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

with observation vector y as

$$\begin{pmatrix} 1 \\ 4 \\ 3 \\ 2 \end{pmatrix}$$

The horizontal coordinate l is assumed to be known exactly in this exercise.

Question 4 With the problem of the bridge deck in the previous question, one can assume that *maximum* deflection occurs (naturally) exactly in the middle, and we conveniently choose the origin of the l-coordinate axis in the middle ($l = 0$), in between the two support points A and B. Present the accordingly simplified model of observation equations.

Answer 4 To find the extremum of the deflection function $d(l)$, we set the first derivative to zero: $\frac{\partial d(l)}{\partial l} = 2lx_1 + x_2 = 0$, which has to occur for $l = 0$, hence $x_2 = 0$. With the later found value for x_1 one can verify that this extremum indeed is a maximum (with positive deflection downward). Now, the second column of the A-matrix together with the given zero value for x_2 can be brought to the left-hand side of the equation. But, as $x_2 = 0$, this cancels all together. The resulting model of observation equations reads

$$E\begin{pmatrix} \underline{d}(l_1) \\ \underline{d}(l_2) \\ \underline{d}(l_3) \\ \underline{d}(l_4) \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 1 & 1 \\ 1 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}$$

Matrix product $A^T A$ is found as

$$A^T A = \begin{pmatrix} 34 & 10 \\ 10 & 4 \end{pmatrix}$$

hence

$$(A^T A)^{-1} = \frac{1}{36} \begin{pmatrix} 4 & -10 \\ -10 & 34 \end{pmatrix}$$

and the least-squares estimates, according to (8.4), are found as

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_3 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 4 & 4 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

and we obtain $\hat{x}_1 = -\frac{2}{3}$ and $\hat{x}_3 = \frac{25}{6}$.