TESTING OF HYPOTHESES

To simplify the calculations, we rewrite the expressions for the various sums of squares as

$$SST = \left\{ \sum_{j=1}^{k} \sum_{i=1}^{n} x_{ij}^2 \right\} - \frac{1}{kn} S^2, \tag{4.38}$$

$$SS(Tr) = \left\{ \frac{1}{n} \sum_{j=1}^{k} S_j^2 \right\} - \frac{1}{kn} S^2, \tag{4.39}$$

and

$$SSE = SST - SS(Tr). \tag{4.40}$$

Here, S_j is the sum of the values in the j'th sample, and S is the total sum of all observations.

So far we assumed that each sample has the same number of observations. Instead, if there are n_j observations in the j'th sample we get

$$SST = \left\{ \sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}^2 \right\} - \frac{1}{N} S^2, \tag{4.41}$$

$$SS(Tr) = \left\{ \sum_{j=1}^{k} \frac{S_j^2}{n_j} \right\} - \frac{1}{N} S^2, \tag{4.42}$$

Where $N = \sum_{j=0}^{k} n_j$. Also, $v_{treat} = k-1$ but $v_{error} = N-k$.

Two-way ANOVA

The previous ANOVA test was concerned only with the task of checking if the means of the populations were the same. The two-way ANOVA procedure extends the test to whether there also are variations *across* the populations. In other words, the population mean for j'th treatment and i'th *block* is expected to be

$$\mu_{ij} = \mu + \varepsilon_j + \gamma_i. \tag{4.43}$$

TESTING OF HYPOTHESES

Thus, ε_j are the treatment effects that vary from sample to sample, and γ_i are called the *block effects* and vary within each sample. Examples might be porosity at four locations where the ε_j represent differences among the locations and γ_i may represent variations with depth across all samples. Again, we test

$$H_0: \varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_k = 0,$$

but now we also consider the second null hypothesis

$$H_0: \gamma_1 = \gamma_2 = \ldots = \gamma_n = 0.$$

To do this we must obtain a quantity, similar to the treatment sum of squares, which measures the variation among the block means. If we let S_i be the total of all values in the i'th block (e.g., depth) and substitute it for S_j, sum over i instead of j, and swap n and k, we find the *block sum of squares* via

$$SSB = \left\{ \frac{1}{k} \sum_{i=1}^{n} S_i^2 \right\} - \frac{1}{kn} S^2. \qquad (4.44)$$

Hence, we compute SST and $SS(Tr)$ as before, SSB as just given, and SSE now becomes

$$SSE = SST - [SS(Tr) + SSB]. \qquad (4.45)$$

Hence, Table 4.7 shows the extended ANOVA table for two-way analysis.

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Treatments	$k-1$	$SS(Tr)$	$MS(Tr) = \frac{SS(Tr)}{k-1}$	$\frac{MS(Tr)}{MSE}$
Blocks	$n-1$	SSB	$MSB = \frac{SSB}{n-1}$	$\frac{MSB}{MSE}$
Error	$(k-1)(n-1)$	SSE	$MSE = \frac{SSE}{(k-1)(n-1)}$	
Total	$kn-1$	SST		

Table 4.7: Two-way table for the analysis of variance (ANOVA).

TESTING OF HYPOTHESES

Example 4–9. We have measured the nickel concentration in a shale at four locations where we have obtained three observations at different depths in the unit. Our data are given in parts per million (ppm), with depth increasing downward in Table 4.8. We want to do a two-way ANOVA to see if the variations among the locations and among observations at the same depth (our "blocks") are similar at the 95% level of confidence. We state

$$H_0: \varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = 0,$$

$$\gamma_1 = \gamma_2 = \gamma_3 = 0.$$

	Loc 1	Loc 2	Loc 3	Loc 4	S_i
Depth 1	71	44	50	67	232
Depth 2	92	51	64	81	288
Depth 3	89	85	72	86	332
S_j	252	180	186	234	852

Table 4.8: Four samples of nickel concentrations from different locations, sorted by depth.

Following the procedure, we compute the total sum of squares to be

$$\sum\sum x^2 = 63{,}414. \tag{4.46}$$

Combined with the sums in Table 4.8 we find

$$SST = 63{,}414 - \frac{1}{12}(852)^2 = 63{,}414 - 60{,}492 = 2922, \tag{4.47}$$

$$SS(Tr) = \frac{1}{3}[252^2 + 180^2 + 186^2 + 234^2] - 60{,}494 = 1260, \tag{4.48}$$

$$SSB = \frac{1}{4}[232^2 + 288^2 + 332^2] - 60{,}492 = 1256, \tag{4.49}$$

$$SSE = 2922 - [1260 + 1256] = 406. \tag{4.50}$$

TESTING OF HYPOTHESES

We construct a two-way ANOVA table presented as Table 4.9.

Source of Variation	ν	SS	MS	F
Treatments	3	1260	420	6.21
Blocks	2	1256	628	9.28
Error	6	406	67.67	
Total	11	2922		

Table 4.9: Two-way ANOVA table resulting from the statistics of the nickel concentrations given in Table 4.8.

With the help of Table A.5 we reach the following decisions:

Treatments (locations): Critical value $F_{0.05,3,6} = 4.76$, so we reject the hypothesis that $\varepsilon_i = 0$.

Blocks (depth): Critical value $F_{0.05,2,6} = 5.14$, so again we reject the hypothesis that all $\gamma_j = 0$.

In other words, we conclude that the average nickel concentration is not the same at the four locations, and that it is not the same at all depths.

TESTING OF HYPOTHESES

4.3 Nonparametric Tests

The last section concluded the examination of standard parametric tests (i.e., the *t*-, *F*-, and χ^2-tests.) We justified using these tests by *either* having large samples and invoking the central limits theorem *or* simply assuming that the distribution we have sampled is approximately normal. Sometimes, however, none of these conditions are met. The two typical situations that can arise are:

1. You have a small sample ($n < 30$) and you *cannot* assume that the population it came from is normal.

2. You have *ordinal* data (which can be ranked, but not operated on numerically).

In those cases we must consider *nonparametric* methods, which make no assumptions about the shape of the data distribution. In particular, nonparametric tests *do not* involve the calculation of distribution parameters, such as the mean and standard deviation.

4.3.1 Sign test for the one-sample mean or median

The nonparametric sign test is a robust alternative to the standard one-sample *t*-test. It can be used when the distribution we have sampled has a continuous symmetrical population. This implies that the probability of getting a data value *less* than the mean is the same as getting one *larger* than the mean: both probabilities equal 0.5. However, if we cannot assume that the population is symmetrical, the test should instead apply to the median value rather than the mean. Since we will be testing whether or not our sample mean (or median) is statistically indistinguishable from a specified hypothetical mean (or median), the procedure relies on properties of the binomial distribution encountered in Chapter 3 and is reminiscent of the simple coin-toss analogy. Values may be less than or larger than the hypothetical mean (median), and the probability of finding *x* values out of *n* values to be less than the median follows directly from (3.70), with $p = 0.5$. To perform the sign test we need to evaluate the cumulative binomial distribution or consult pre-tabulated distributions.

TESTING OF HYPOTHESES

Example 4–10. The following data constitute a random sample of 15 measurements of salinity content (in ppt):

97.5, 95.2, 97.3, 96.0, 96.8, 100.3, 97.4, 95.3, 93.2, 99.1, 96.1, 97.6, 98.2, 98.5, 94.9

We will use the one-sample sign test to test the null hypothesis, H_0: $\tilde{\mu} \geq 98.5$ against the alternative hypothesis, H_1: $\tilde{\mu} < 98.5$ at the $\alpha = 0.01$ level of significance. Because of the inequality we have a one-sided test. We replace all values greater than 98.5 with a plus sign and all values less than 98.5 with a minus sign. Values that equal 98.5 exactly are discarded; in our case we lose one value, thus $n = 14$, resulting in the following series:

-----+---+----

We find $x = 2$ values (represented by the two plus-signs) larger than the hypothetical median. The probability of finding $x \leq 2$ is given by the binomial distribution (3.70) by adding up the probabilities for $x = 0$, $x = 1$, and $x = 2$. We find

$$P = P_{14,0.5}(0) + P_{14,0.5}(1) + P_{14,0.5}(2) = C_{14,0.5}(2) = \sum_{x=0}^{2} P_{14,0.5}(x), \tag{4.51}$$

which evaluates as

$$P = \binom{14}{0}\frac{1}{2}^{14} + \binom{14}{1}\frac{1}{2}^{14} + \binom{14}{2}\frac{1}{2}^{14} = 0.00006 + 0.0009 + 0.0056 \approx 0.0065. \tag{4.52}$$

Since 0.0065 is less than 0.01, we must reject H_0. We conclude that the null hypothesis must be rejected as the data suggest that the median salinity from the sampled region is less than 98.5 ppt. Note that in this test we did not compute a critical value for x but compared the probability for the observed case with a specified probability α.

TESTING OF HYPOTHESES

When both np and $n(1-p)$ are greater than 5 (here they are both equal to 7) we are allowed to use the normal approximation to the binomial distribution. Per (3.87), the sign test may then be based on the statistic

$$z = \frac{x - np}{\sqrt{np(1-p)}}, \tag{4.53}$$

which in our situation (with $p = 0.5$) simplifies to

$$z = \frac{2x - n}{\sqrt{n}}. \tag{4.54}$$

We may now simply compare the observed z statistic with the chosen $z_{\alpha/2}$ critical value as in the standard parametric case (or z_α for a one-sided test like the present case). Here, $z_\alpha = -2.326$ while observed $z = -2.676$ and we again must reject the null hypothesis.

4.3.2 Mann-Whitney test (*U*-test)

This test is a nonparametric alternative to the two-sample Student's *t*-test. It also goes by the names *Wilcoxon* test and the *U* -test. The Mann-Whitney test is performed by combining the two data sets we want to compare, sorting the combined set into ascending order, and assigning each point a *rank*: the smallest value is given rank = 1, while the largest observation is ranked $n_1 + n_2$. Should some of the observations be identical one assigns the average rank to all tied values. E.g., if the 7th and 8th sorted values were identical, we would assign to each the rank of 7.5. The idea here is that if the samples consist of random drawings from the same population (i.e., when H_0 is true) then we would expect the ranks for both samples to be scattered more-or-less uniformly throughout the sequence. This would be true regardless of the distribution that characterizes the population.

TESTING OF HYPOTHESES

After sorting the data we add up the ranks for each data set separately into *rank sums*, which we denote S_1 and S_2. The sum of S_1+S_2 must obviously equal the sum of the first (n_1+n_2) integers, which is

$$\frac{1}{2}(n_1+n_2)(n_1+n_2+1). \tag{4.55}$$

Many early rank sum tests were based on S_1 or S_2 but now it is customary to use the statistic U, defined as $U = \min(U_1,U_2)$, i.e., the smallest of U_1 and U_2, with

$$U_1 = S_1 - \frac{1}{2}n_1(n_1+1) \tag{4.56}$$

and

$$U_2 = S_2 - \frac{1}{2}n_2(n_2+1). \tag{4.57}$$

This statistic can range from 0 to $n_1 \cdot n_2$ and its sampling distribution is symmetrical about $n_1 \cdot n_2/2$. The test, then, consists of these steps:

1. Compute the U -statistic using the smallest value of (4.56) and (4.57).
2. Given the sample sizes and the desired level of significance α, evaluate critical $U_{\alpha,n1,n2}$.
3. Compare the calculated U statistic to the critical $U_{\alpha,n1,n2}$ and reject H_0 if our U is *less than* the critical value.

Note that for the U-test, H_0 is rejected when our U -value is *less* than and not larger than the critical value, as is common for most other tests we have discussed.

TESTING OF HYPOTHESES

Example 4–11. We want to compare the grain size of sand obtained from two different locations on the moon on the basis of measurements of grain diameters (in mm), as follows:

Location 1	0.37	0.70	0.75	0.30	0.45	0.16	0.62	0.73	0.33	$n_1 = 9$
Location 2	0.86	0.55	0.80	0.42	0.97	0.84	0.24	0.51	0.92	0.69	$n_2 = 10$

We do not know what type of distribution that grain sizes of sand on the moon might follow, so we choose the *U* -test to see if the mean grain sizes differ between the two samples. Computing the sample means gives $\bar{x}_1 = 0.49$ and $\bar{x}_2 = 0.68$. If we wanted to use the *t*-test we would have to assume that the underlying distributions are normal, since the samples are small. The *U* -test requires no such assumptions. We start by arranging the data jointly into ascending order and keep track of which location each point originated from (Table 4.10).

We first evaluate the rank sum for location 1, giving $S_1 = 69$, from which it follows that

$$S_2 = \frac{19 \cdot 20}{2} - S_1 = 190 - 69 = 121. \tag{4.58}$$

We now form the null hypothesis $H_0: \mu_1 = \mu_2$, with $H_1: \mu_1 \neq \mu_2$, and state the level of significance $\alpha = 0.05$. Table A.10 has critical values for *U* and we find $U_{\alpha,9,10} = 20$. We will reject the null hypothesis if *U* is ≤ 20. From S_1 and S_2 we find

$$U_1 = 69 - \frac{9 \cdot 10}{2} = 24, \tag{4.59}$$

$$U_2 = 121 - \frac{10 \cdot 11}{2} = 66, \tag{4.60}$$

and hence $U = \min(24,66) = 24$. This is larger than the critical value of 20, suggesting we cannot reject the null hypothesis. In other words, the observed difference in mean grain size at the two locations is not statistically significant at the 95% level of confidence.

TESTING OF HYPOTHESES

For large samples (n_1, $n_2 > 30$) the procedure again simplifies and it can be shown that the mean and standard deviation of the U sampling distribution approach

$$\mu_U = \frac{n_1 n_2}{2}, \quad \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}, \tag{4.61}$$

provided there are no tied ranks. We could then evaluate standard z-scores as $z = (U - \mu_U)/\sigma_U$ and use the familiar critical values $\pm z_{\alpha/2}$ from Table A.2.

4.3.3 Comparing distributions: The Kolmogorov-Smirnov test

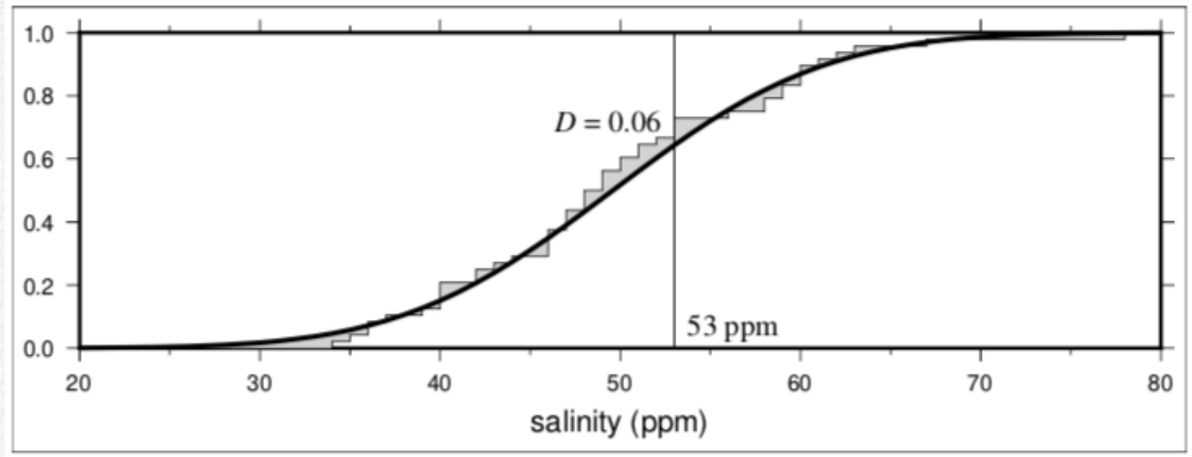

Figure 4.7: Solid line is a cumulative normal distribution for μ = 49.59 ppm and σ = 9.27 ppm. The stair-case curve is the observed cumulative distribution which has its maximum difference, D, from the theoretical curve at the 53 ppm point.

Another very useful nonparametric method is the Kolmogorov-Smirnov test (or K-S for short). It is a test for goodness of fit or *shape* and is often used instead of the χ^2-test. We may use it to test the null hypothesis that two distributions have the same probability density function (i.e., the same shape). A big advantage of the K-S test over the χ^2-test is that one does not have to bin the data, which is an arbitrary procedure anyway (how do you select bin size and why?). In the K-S test we convert the data distribution to a cumulative distribution $C(x)$. Clearly, $C(x)$ then gives the fraction of data points to the "left" of x. While different data sets will in general have different distributions, all cumulative distributions agree at the smallest x ($C(x) \equiv 0$) and at the largest x ($C(x) \equiv 1$). Thus, it is the behavior *between* these points that sets distributions apart (e.g., Figure 4.7). There is of course an infinite number of ways to measure the overall difference between two cumulative distributions: we could look at the absolute value of the area between the curves, the mean square difference, etc. The K-S statistic chooses a simple measure: It determines the maximum absolute difference between the two cumulative curves. Thus, when comparing two cumulative distributions $C_1(x)$ and $C_2(x)$ our K-S statistic becomes

$$D = \max_{-\infty < x < \infty} |C_1(x) - C_2(x)|. \tag{4.61}$$

TESTING OF HYPOTHESES

Note that C_2 may be another data-derived cumulative distribution (the test is then a two-sample test) or a theoretical cumulative probability function like the cumulative normal distribution (we call this case a one-sample test). The distribution of the K-S statistic itself can be calculated under the assumption that C_1 and C_2 are drawn from the same distribution (i.e., H_0), thus providing critical values for D. For the one-sample scenario there are two different cases to consider:

1. C_2 is a *known* cumulative distribution function, i.e., its parameters are prescribed by the null hypothesis.

2. C_2 is an *unknown* cumulative distribution function, i.e., its parameters must first be computed from C_1.

The former case is the problem studied by Kolmogorov and Smirnov. The latter case, however, clearly reduces the degrees of freedom and the standard K-S critical values are overestimated. A different set of critical values (called the *Lilliefors* critical values) has been developed for the normal distribution. Hence, when we need to compute the mean and standard deviation first we say we are performing a *Lilliefors test* rather than a Kolmogorov-Smirnov test.

We will use this test on the salinity measurements we looked at previously (Table 4.2). We sort the salinity measurements, convert them to a cumulative distribution (e.g., C_1), and plot the cumulative function on the same graph as that of a normal cumulative distribution with the same mean and standard deviation (e.g., C_2). Inspecting Figure 4.7 we find the maximum absolute difference to occur at the 53 ppt observation. The D estimate is 0.701 − 0.641 = 0.06. Based on a significance level of α = 0.05 and n = 48, the critical Lilliefors value for a two-sided test is found in Table A.13 to be ~ 0.128, which is much larger than observed. Hence we cannot reject the null hypothesis that the samples were collected from a normally distributed population. In this example, both the K-S and χ^2 tests reached the same conclusion.

4.3.4 Spearman's rank correlation

Finally, we will look at nonparametric correlation called Spearman's *rank correlation*, denoted by r_s. The rank correlation is carried out by ranking the x_i's and y_i's *separately*, then computing the standard correlation coefficient (i.e., 4.29) using the ranks *in lieu* of the data values. Let u_i be the rank of the i'th pair's x-value and v_i be the rank of the i'th pair's y-value. Then, Spearman's rank correlation depends on the covariance and variances of the ranks:

$$r_s = \frac{s_{uv}}{s_u s_v} = \frac{n\sum_{i=1}^{n} u_i v_i - \left(\sum_{i=1}^{n} u_i\right)\left(\sum_{i=1}^{n} v_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} u_i^2 - \left(\sum_{i=1}^{n} u_i\right)^2\right]\left[n\sum_{i=1}^{n} v_i^2 - \left(\sum_{i=1}^{n} v_i\right)^2\right]}}. \qquad (4.63)$$

TESTING OF HYPOTHESES

If there are runs of tied ranks then we assign those points their *average* rank. Fortunately, for situations where there are no ties (4.63) simplifies greatly to

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}, \tag{4.64}$$

where $d_i = u_i - v_i$ is the difference in rank for each (x_i, y_i) pair. In the case where the null hypothesis H_0: ρ = 0 is true, the sampling distribution of r is approximately normal and has zero mean ($\mu = 0$) and standard deviation $\sigma = 1/\sqrt{n-1}$. We could therefore base our statistics on

$$z = \frac{r_s - \mu}{\sigma} = \frac{r_s - 0}{1/\sqrt{n-1}} = r_s\sqrt{n-1} \tag{4.65}$$

and compare this observed z-value to critical $z_{\alpha/2}$ values. However, it turns out that a better approximation is the one given by (4.30), which we utilized when testing the standard correlation coefficient. Even so, for small data sets ($n < 20$) either approximation deviates from the true distribution and hence special tables are required (see Table A.15).

A comparison between the standard correlation and the Spearman's rank correlation reveals some interesting differences:

- Spearman's rank correlation is more tolerant of outliers since only their outlying *ranks* and not actual data values enter into the calculation.
- While the standard correlation measures the degree of *linear* correlation between x and y, Spearman's rank correlation measures the degree of *monotonicity* of the two rank series. Any data set whose ranks u and v vary monotonically will yield $r_s = \pm 1$, even if they do not form a linear trend.

In most other situations the two correlations will be similar.

TESTING OF HYPOTHESES

Ranking the dice data discussed earlier (Table 4.3) gives the values listed in Table 4.11. Using (4.64) we find $r_s = 0.65$ (surprisingly similar to the $r = 0.66$ we found using (4.29)), while the exact equation (4.63) yields $r_s = 0.6325$, which may be close enough for government work. For the simplified value the z-statistic from (4.65) becomes $z = 1.3$, which is well inside the 95% confidence limits ($z_{0.025} = \pm 1.96$) for a normal distribution. Likewise, using (4.30) we find $t = 1.41$ with critical $t_{0.025,3} = 3.18$. Hence, in either case we again arrive at the same conclusion that we cannot reject H_0. However, for such a small data set the approximations usually are quite poor; Table A.15 states the critical correlation is 1, meaning it would take a perfect nonparametric correlation to reject H_0.

Red (x)	**Rank** x	**Green** (y)	**Rank** y	**d**
4	3.5	5	4	0.5
2	1.5	2	2	0.5
4	3.5	6	5	1.5
2	1.5	1	1	-0.5
6	5	4	3	-2

Table 4.11: Evaluating the differences in ranks among $x - y$ pairs obtained by rolling red and green dice. Notice there are two groups of x-values with tied ranks but none among the y-values.

In summary, there are numerous tests, both parametric and nonparametric, that can be applied to our data, and there are many others not covered in these notes. However, the ones presented here are the most common hypothesis tests that all scientists should be aware of. A simple guide to their use is given in Figure 4.8.

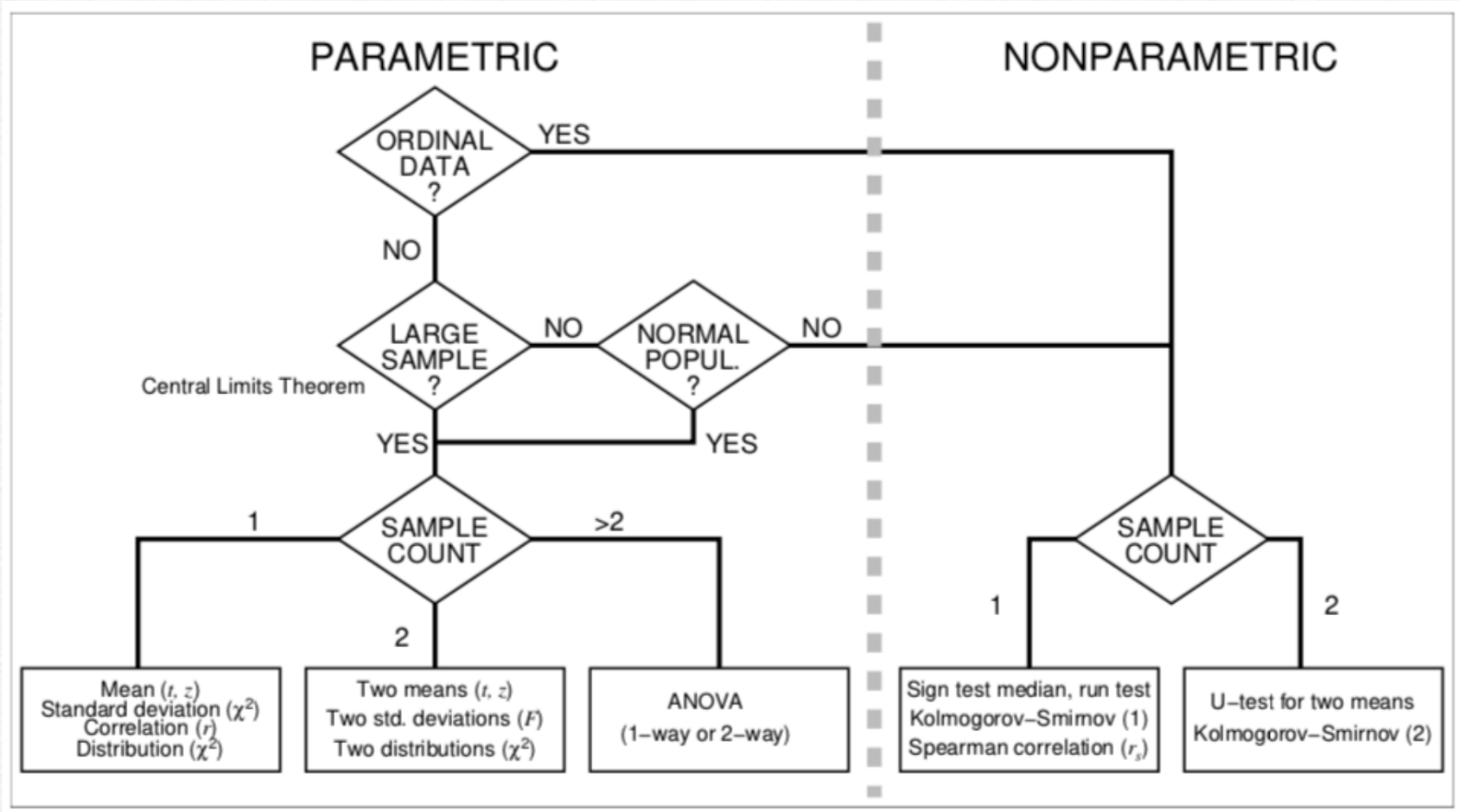

Figure 4.8: Simple decision chart for selecting standard parametric or nonparametric tests. The *run test* is a specific application of the sign test and will be discussed in Chapter 7.

LINEAR (MATRIX) ALGEBRA

5.9 Solution of Simultaneous Linear Equations

A system of four simultaneous linear equations in four unknowns x_1, x_2, x_3, x_4 can be written

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 &= b_4 \end{aligned} \tag{5.77}$$

or, in matrix form,

$$\mathbf{Ax=b}, \tag{5.78}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \tag{5.79}$$

is called the coefficient matrix,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{5.80}$$

is the unknown vector, and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \tag{5.81}$$

is the right hand side (i.e., the observations). Premultiplying both sides by A^{-1} yields

$$\mathbf{A^{-1}Ax = A^{-1}b}, \tag{5.82}$$

LINEAR (MATRIX) ALGEBRA

hence

$$\mathbf{I}\,\mathbf{x}=\mathbf{x}=\mathbf{A}^{-1}\mathbf{b} \tag{5.83}$$

gives the solution for values of x_1, x_2, x_3, x_4 which solve the system. For simplicity, the following example solves for two simultaneous equations only. Consider two equations in two unknowns (e.g., equations of lines in the $x_1 - x_2$ plane):

$$\begin{aligned} 5x_1 + 7x_2 &= 19 \\ 3x_1 - 2x_2 &= -1 \end{aligned} \tag{5.84}$$

In matrix form this system translates to

$$\begin{bmatrix} 5 & 7 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ -1 \end{bmatrix} \tag{5.85}$$

or

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b}. \tag{5.86}$$

To solve this matrix equation we need the inverse of A, which is simply

$$\mathbf{A}^{-1} = \frac{1}{-10-21} \begin{bmatrix} -2 & -7 \\ -3 & 5 \end{bmatrix} = \begin{bmatrix} \frac{2}{31} & \frac{7}{31} \\ \frac{3}{31} & \frac{-5}{31} \end{bmatrix} \tag{5.87}$$

Then, $\mathbf{x} = \mathbf{A}^{-1}\cdot\mathbf{b}$, where

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} \frac{2}{31} & \frac{7}{31} \\ \frac{3}{31} & \frac{-5}{31} \end{bmatrix} \begin{bmatrix} 19 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{38}{31} - \frac{7}{31} \\ \frac{57}{31} + \frac{5}{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{5.88}$$

So, the values $x_1 = 1$ and $x_2 = 2$ solve the above system, or

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{5.89}$$

While this approach may seem burdensome, it is good because it is extremely general and allows for a straight- forward solution to very large systems. However, it is true that direct (elimination) methods to the solution are in fact quicker for fully populated matrices:

1. A solution using the inverse matrix approach involves n^3 multiplications for the inversion and n^2m more multiplications to finish the solution, where n is the number of equations per set, and m is the number of sets of equations (each of the same form but different b vector). The total number of multiplications is n^3+n^2m.

2. A solution by directly solving the linear equations involves $n^3/3+n^2m$ multiplications.

Hence, while the matrix form is easy to handle, one should not necessarily always use it blindly. We will consider many situations for which matrix solutions are ideal. For sparse or symmetrical matrices, the above relationships may not hold.

5.9.1 Simple regression and curve fitting

Whereas an interpolant fits each data point exactly, it is frequently advantageous to produce a smoothed fit to the data — not exactly fitting each point, but producing a "best" fit. A popular (and convenient) method for producing such fits is known as the *method of least squares*.

The method of least squares produces a fit of a specified (usually continuous) basis to a set of data points which minimizes the sum of the squared misfit (error) between the fitted curve and the data. The misfit can be measured vertically, as in Figure 5.4. This *regression* of y on x is the most commonly used method. Less common methods (i.e., more work involved) is the regression of x on y and even orthogonal regression (which we will return to later; see Figure 5.5).

Consider fitting a single "best" linear curve to n data points. This can be a scatter plot of $x(t)$, $d(t)$ plotted at similar values of t, or a simple $d = f(x)$ relationship. At any rate, d (our data) are considered a function of x (which may be a spatial coordinate or time). We wish to fit a line of the form

$$d(x) = m_1 + m_2(x - x_0) \tag{5.90}$$

LINEAR (MATRIX) ALGEBRA

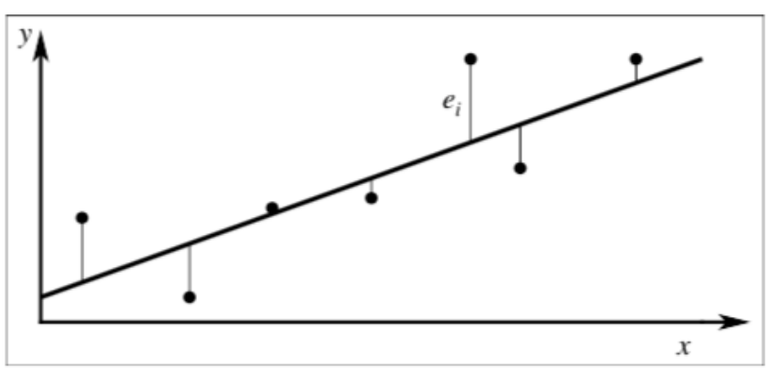

Figure 5.4: Graphical representation of the regression errors used in least-squares procedures. We measure misfit vertically in the *y*-direction from data point to regression curve.

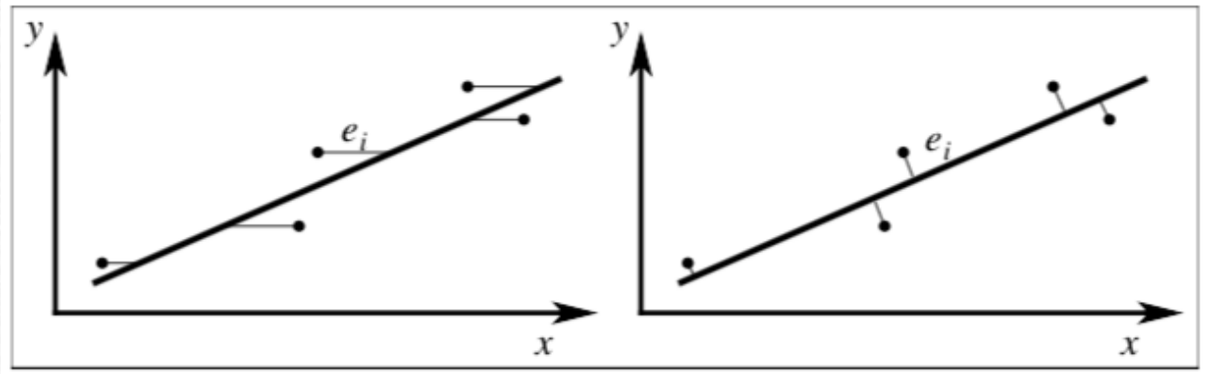

Figure 5.5: Two other regression methods: regressing *x* on *y* and orthogonal regression. Here we measure misfits horizontally from data point to regression line or orthogonally onto the regression line, respectively.

and must therefore determine values for the model coefficients m_1 and m_2 that produce a line that minimizes the sum of the squared misfits (here, x_0 is a constant specified beforehand). In other words,

$$\text{minimize} \sum_{i=1}^{n} \left[(d_{\text{computed}}(x_i) - d_{\text{observed}}(x_i) \right]^2 \tag{5.91}$$

Ideally, for each observation d_i at location x_i we should have

$$\begin{aligned} m_1 + m_2(x_1 - x_0) &= d_1 \\ m_1 + m_2(x_2 - x_0) &= d_2 \\ m_1 + m_2(x_3 - x_0) &= d_3 \\ &\vdots \\ m_1 + m_2(x_n - x_0) &= d_n \end{aligned} \tag{5.92}$$

LINEAR (MATRIX) ALGEBRA

There are many more equations (n — one for each observed value of d) than unknowns (2: m_1 and m_2). Such a system is *overdetermined* and there exists no unique solution (unless all the d_i's do lie exactly on a single line, in which case any two equations will uniquely determine m_1 and m_2). In matrix form,

$$\begin{bmatrix} 1 & (x_1 - x_0) \\ 1 & (x_2 - x_0) \\ \vdots & \vdots \\ 1 & (x_n - x_0) \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \tag{5.93}$$

i.e., **G · m = d**. Here, **G** represents how the data **d** are related to the model **m** and is often called the *design matrix*. However, since **G** is a not square it has no inverse, hence the equation cannot be inverted and solved as is. Consider instead the *misfit*, e_i, at each point:

$$\begin{aligned} m_1 + m_2(x_1 - x_0) - d_1 &= e_1 \\ m_1 + m_2(x_2 - x_0) - d_2 &= e_2 \\ &\vdots \\ m_1 + m_2(x_n - x_0) - d_n &= e_n \end{aligned} \tag{5.94}$$

We wish to determine the values for m_1 and m_2 that minimize

$$E(m_1, m_2) = \sum_{i=1}^{n} e_i^2 = \mathbf{e}^T \mathbf{e}, \tag{5.95}$$

where $\mathbf{e}^T = (e_1, e_2, \ldots, e_n)$ is the *misfit vector*. This condition will minimize the variance of the residuals about the regression line and give the desired least-squares fit. Thus, $E(m_1, m_2)$ and the minimum of this function (with respect to the two unknown coefficients) can be determined using simple differential calculus where, at the desired minimum, we require

$$\frac{\partial E(m_1, m_2)}{\partial m_1} = \frac{\partial E(m_1, m_2)}{\partial m_2} = 0. \tag{5.96}$$

LINEAR (MATRIX) ALGEBRA

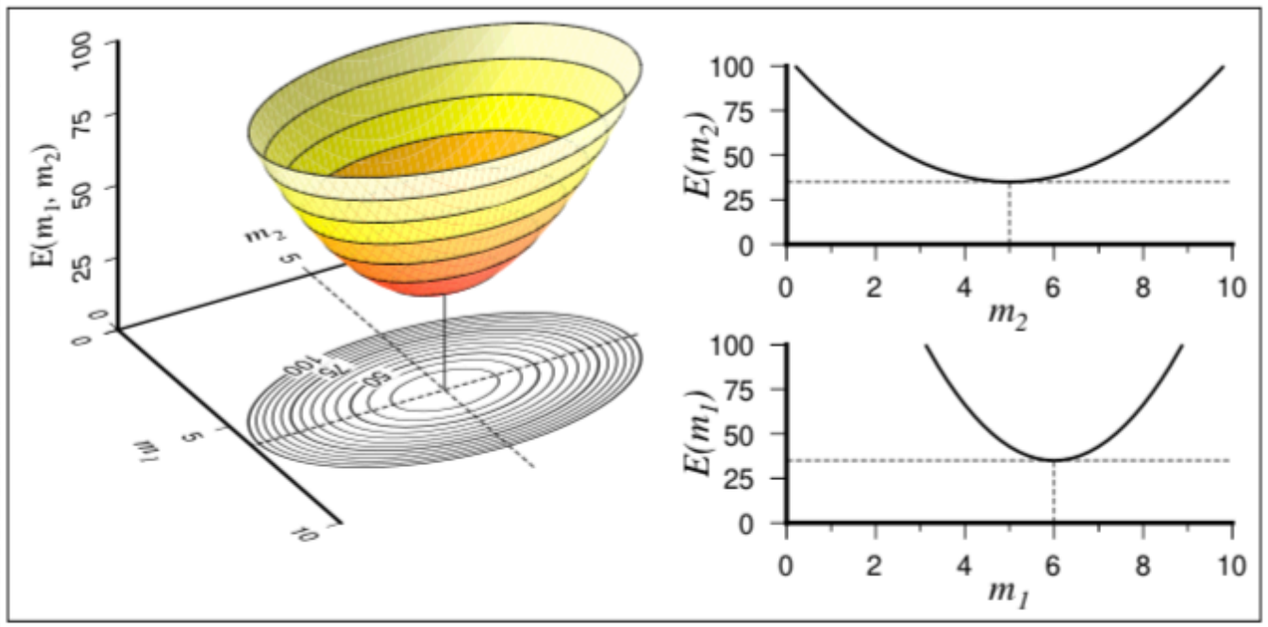

Figure 5.6: (left) The solution we seek minimizes the misfit function $E(\mathbf{m}) = E(m_1,m_2)$, which portrays a surface in 3-D. Because of the functional (quadratic) form of E we are guaranteed a unique global minimum. (right) Two orthogonal cross-sections of E along the axes m_1 and m_2. We seek the solutions for these two parameters so that the respective slopes in E are zero simultaneously.

Thus, the *slopes* of the misfit function with respect to each parameter must be zero (see Figure 5.6). We find

$$\frac{\partial E}{\partial m_1} = \frac{\partial}{\partial m_1}\left(\sum_{i=1}^{n} e_i^2\right) = \frac{\partial}{\partial m_1}\left\{\sum_{i=1}^{n}\left[m_1 + m_2(x_i - x_0) - d_i\right]^2\right\}$$

$$= 2\sum_{i=1}^{n}\left[m_1 + m_2(x_i - x_0) - d_i\right] = 0 \qquad (5.97)$$

$$\frac{\partial E}{\partial m_2} = \frac{\partial}{\partial m_2}\left(\sum_{i=1}^{n} e_i^2\right) = \frac{\partial}{\partial m_2}\left\{\sum_{i=1}^{n}\left[m_1 + m_2(x_i - x_0) - d_i\right]^2\right\}$$

$$= 2\sum_{i=1}^{n}\left[m_1 + m_2(x_i - x_0) - d_i\right](x_i - x_0) = 0. \qquad (5.98)$$

These two equations can now be expanded into their individual terms, forming what are known as the *normal equations*.

LINEAR (MATRIX) ALGEBRA

This system of two equations with two unknowns can be uniquely solved. Rearranging, we find

$$nm_1 + m_2 \sum_{i=1}^{n} (x_i - x_0) = \sum_{i=1}^{n} d_i, \tag{5.99}$$

$$m_1 \sum_{i=1}^{n} (x_i - x_0) + m_2 \sum_{i=1}^{n} (x_i - x_0)^2 = \sum_{i=1}^{n} d_i (x_i - x_0). \tag{5.100}$$

Notice that all sums involved are known values that add to simple constants. Specifically, we must compute the sums

$$S_y = \sum_{i=1}^{n} d_i,\ S_{xy} = \sum_{i=1}^{n} d_i (x_i - x_0), S_x = \sum_{i=1}^{n} (x_i - x_0),\ \text{and}\ S_{xx} = \sum_{i=1}^{n} (x_i - x_0)^2. \tag{5.101}$$

Substituting these symbols into (5.99) and (5.100), we obtain

$$nm_1 + m_2 S_x = S_y \tag{5.102}$$

$$m_1 S_x + m_2 S_{xx} = S_{xy} \tag{5.103}$$

Solving for the intercept yields

$$m_1 = \frac{1}{n} S_y - \frac{m_2}{n} S_x. \tag{5.104}$$

We substitute m_1 into (5.103) and find

$$\left[\frac{1}{n} S_y - \frac{m_2}{n} S_x \right] S_x + m_2 S_{xx} = S_{xy}. \tag{5.105}$$

Now solve for m_2:

$$\frac{1}{n} S_y S_x - \frac{m_2}{n} S_x^2 + m_2 S_{xx} = S_{xy}, \tag{5.106}$$

$$m_2 \left(S_{xx} - \frac{1}{n} S_x^2 \right) = S_{xy} - \frac{1}{n} S_y S_x. \tag{5.107}$$

LINEAR (MATRIX) ALGEBRA

Finally,

$$m_2 = \left(S_{xy} - \frac{1}{n}S_y S_x\right) / \left(S_{xx} - \frac{1}{n}S_x^2\right) = \frac{nS_{xy} - S_x S_y}{nS_{xx} - S_x^2}, \tag{5.108}$$

and we substitute m_2 into (5.104) to find

$$m_1 = \frac{S_{xx}S_y - S_x S_{xy}}{nS_{xx} - S_x^2}. \tag{5.109}$$

In matrix form the normal equations are

$$\begin{bmatrix} n & \sum_{i=1}^{n}(x_i - x_0) \\ \sum_{i=1}^{n}(x_i - x_0) & \sum_{i=1}^{n}(x_i - x_0)^2 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} d_i \\ \sum_{i=1}^{n} d_i(x_i - x_0) \end{bmatrix} \tag{5.110}$$

which may be simplified to

$$\begin{bmatrix} n & S_x \\ S_x & S_{xx} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} S_y \\ S_{xy} \end{bmatrix} \tag{5.111}$$

Therefore, Nm = v, and since N is square, symmetric and of full rank, this equation is solved in the standard manner:

$$\mathbf{N}^{-1}\mathbf{N}\mathbf{m} = \mathbf{m} = \mathbf{N}^{-1}\mathbf{v}. \tag{5.112}$$

This problem was simple enough (2 × 2) to solve for m_1 and m_2 by brute force. For larger systems, this approach becomes impractical and instead a matrix solution to the rectangular G · m = d equation must be sought. We will next look at the general linear least-squares problem and find a solution in matrix notation.

5.9.2 General linear least squares method, version 1

We have looked at a few special cases where we have sought to fit a model to data in a least-squares sense. Fitting a straight line to the $x - d$ points was a very simple example of this technique. We will now look at the more general problem of finding the coefficients for *any* linear combination of a chosen set of basis functions that fits a data set in a least squares sense. There are numerous situations where this is needed; some are listed in Table 5.1.

LINEAR (MATRIX) ALGEBRA

Situation	Model	Data
Curve Fitting	Coefficients of polynomials (x^m), sin, cos, etc.	Points in $x-y$ plane
Gravity modeling	Densities of subsurface polygons	Gravity observations
Hypocenter location	Small perturbations to hypocenter location	Seismic arrival times

Table 5.1: Examples of situations where linear least squares solutions are used.

Polynomial basis	Fourier sine basis
$g_1 = x^0$	$g_1 = \sin(2\pi x/T)$
$g_2 = x^1$	$g_2 = \sin(4\pi x/T)$
$g_3 = x^2$	$g_3 = \sin(6\pi x/T)$
$\vdots$	$\vdots$
$g_m = x^{m-1}$	$g_m = \sin(2m\pi x/T)$

Table 5.2: Examples of basis functions used for modeling of data.

While the basis functions in these cases are all vastly different, they are all used in linear combinations to fit the observed data. We will therefore take time to investigate how such a problem is set up, and how the setup can be simplified with matrix algebra. Some typical basis functions are given in Table 5.2.

Consider the least squares fitting of any continuous basis of the form

$$g_1(x),\ g_2(x),\ g_3(x),\ \cdots,\ g_m(x). \tag{5.113}$$

For example, we desire to fit a model with m terms

$$d(x) = m_1 g_1(x) + m_2 g_2(x) + \cdots + m_m g_m(x) \tag{5.114}$$

to a data set of n data points, where $n > m$, by minimizing $E(\mathbf{m})$ given by

$$E(\mathbf{m}) = E(m_1, m_2, \cdots, m_m) = \sum_{i=1}^{n} (e_i)^2 = \sum_{i=1}^{n} (m_1 g_1(x_i) + m_2 g_2(x_i) + \cdots + m_m g_m(x_i) - d_i)^2, \tag{5.115}$$

or simply

$$E(\mathbf{m}) = \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots + m_m g_{im} - d_i)^2, \tag{5.116}$$

where d_i is the observed value and g_{ij} is the jth basis function, evaluated at the location (or time) x_i. In other words, $g_{ij} = g_j(x_i)$.

LINEAR (MATRIX) ALGEBRA

There are *four* concepts of vital importance in a general linear least squares modeling problem:

1. The *observed data*, $(x_i,d_i),i = 1,n$, where n is the number of observations. These are all known quantities.
2. The *general linear model* (linear in m=m_j, j=1, m, with m unknown parameters), given by (5.114).
3. The m chosen *basis functions*, $g_j(x)$, $j = 1, m$. We can evaluate these for any x.
4. The *least squares misfit criteria*, given by (5.116).

We can write a linear system of equations for the misfit at each data point:

$$\begin{aligned} m_1 g_{11} + m_2 g_{12} + \cdots + m_m g_{1m} - d_1 &= e_1 \\ m_1 g_{21} + m_2 g_{22} + \cdots + m_m g_{2m} - d_2 &= e_2 \\ &\vdots \\ m_1 g_{n1} + m_2 g_{n2} + \cdots + m_m g_{nm} - d_n &= e_n \end{aligned} \tag{5.117}$$

To minimize E, we require

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = 0, j = 1, m. \tag{5.118}$$

Considering the first term (case j = 1), we see

$$\begin{aligned} \frac{\partial E(\mathbf{m})}{\partial m_1} &= \frac{\partial}{\partial m_1} \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots m_m g_{im} - d_i)^2 \\ &= 2 \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots m_m g_{im} - d_i) g_{i1} = 0, \end{aligned} \tag{5.119}$$

while for the second term (case j = 2), we find

$$\begin{aligned} \frac{\partial E(\mathbf{m})}{\partial m_2} &= \frac{\partial}{\partial m_2} \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots m_m g_{im} - d_i)^2 \\ &= 2 \sum_{i=1}^{n} (m_1 g_{i1} + m_2 g_{i2} + \cdots m_m g_{im} - d_i) g_{i2} = 0. \end{aligned} \tag{5.120}$$

LINEAR (MATRIX) ALGEBRA

Consequently, for the *j*'th parameter,

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = 2\sum_{i=1}^{n}(m_1 g_{i1} + m_2 g_{i2} + \cdots m_m g_{im} - d_i)g_{ij} = 0. \tag{5.121}$$

Rearranging these normal equations gives the square $m \times m$ system

$$\begin{aligned}
m_1\sum_{i=1}^{n} g_{i1}^2 + m_2\sum_{i=1}^{n} g_{i2}g_{i1} + \cdots + m_m\sum_{i=1}^{n} g_{im}g_{i1} &= \sum_{i=1}^{n} d_i g_{i1} \\
m_1\sum_{i=1}^{n} g_{i1}g_{i2} + m_2\sum_{i=1}^{n} g_{i2}^2 + \cdots + m_m\sum_{i=1}^{n} g_{im}g_{i2} &= \sum_{i=1}^{n} d_i g_{i2} \\
&\vdots \\
m_1\sum_{i=1}^{n} g_{i1}g_{im} + m_2\sum_{i=1}^{n} g_{i2}g_{im} + \cdots + m_m\sum_{i=1}^{n} g_{im}^2 &= \sum_{i=1}^{n} d_i g_{im}
\end{aligned} \tag{5.122}$$

or equivalently,

$$m_1\sum_{i=1}^{n} g_{i1}g_{ij} + m_2\sum_{i=1}^{n} g_{i2}g_{ij} + \cdots + m_m\sum_{i=1}^{n} g_{im}g_{ij} = \sum_{i=1}^{n} d_i g_{ij}, j = 1, m. \tag{5.123}$$

This setup provides a *closed system* of *m* normal equations. In matrix form,

$$\begin{bmatrix}
\sum_{i=1}^{n} g_{i1}^2 & \sum_{i=1}^{n} g_{i2}g_{i1} & \cdots & \sum_{i=1}^{n} g_{im}g_{i1} \\
\sum_{i=1}^{n} g_{i1}g_{i2} & \sum_{i=1}^{n} g_{i2}^2 & \cdots & \sum_{i=1}^{n} g_{im}g_{i2} \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{i=1}^{n} g_{i1}g_{im} & \sum_{i=1}^{n} g_{i2}g_{im} & \cdots & \sum_{i=1}^{n} g_{im}^2
\end{bmatrix}
\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_m \end{bmatrix} =
\begin{bmatrix} \sum_{i=1}^{n} d_i g_{i1} \\ \sum_{i=1}^{n} d_i g_{i2} \\ \vdots \\ \sum_{i=1}^{n} d_i g_{im} \end{bmatrix} \tag{5.124}$$

Hence, we simply have

$$\mathbf{N}\cdot\mathbf{m}=\mathbf{v}, \tag{5.125}$$

where N is the (known) coefficient matrix, m the vector with the unknowns m_j, and v contains weighted sums of known (observed or computable) quantities. Solving for the m vector (since N is square, symmetric and of full rank) yields

$$\mathbf{N}^{-1}\cdot\mathbf{N}\cdot\mathbf{m}=\mathbf{m}=\mathbf{N}^{-1}\cdot\mathbf{v}. \tag{5.126}$$

The resulting m_j values are the ones which satisfy (5.118) and therefore the same ones, when combined with the chosen basis, that produce the "best" fit to the *n* data points such that (5.116) is minimized.

5.9.3 General linear least squares method, version 2

We will now look at a simpler approach to the same problem using matrix algebra. We have **e = G · m – d**, or

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1m} \\ g_{21} & g_{22} & \cdots & g_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nm} \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_m \end{bmatrix} - \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \tag{5.127}$$

We note that each column vector of **G** is simply a single basis function evaluated at all our observation points. In fact, we could write **G** as

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_m \end{bmatrix}, \tag{5.128}$$

where

$$\mathbf{g}_j = \begin{bmatrix} g_j(x_1) & g_j(x_2) & \cdots & g_j(x_n) \end{bmatrix}^T \tag{5.129}$$

We wish to find the m_j values that minimize $E = \mathbf{e}^T\mathbf{e}$. Minimizing the misfit with respect to the unknown model parameters m means we must solve the *m* linear equations that result from setting all partial derivatives of *E* to zero (i.e., 5.118).

LINEAR (MATRIX) ALGEBRA

Using matrix algebra, we express the *predicted* solution as $\hat{\boldsymbol{d}}$ = **G · m**. We may now express the misfit between model and observations as **e** = $\hat{\boldsymbol{d}}$ **– d = G · m – d** and use this expression to evaluate the misfit as

$$E(\mathbf{m}) = \sum_{i=1}^{n} (e_i)^2 = \mathbf{e}^T \cdot \mathbf{e} = \left(\hat{\mathbf{d}} - \mathbf{d}\right)^T \cdot \left(\hat{\mathbf{d}} - \mathbf{d}\right) = (\mathbf{G}\cdot\mathbf{m} - \mathbf{d})^T \cdot (\mathbf{G}\cdot\mathbf{m} - \mathbf{d}). \tag{5.130}$$

Expanding terms, we find

$$E(\mathbf{m}) = \left(\mathbf{m}^T\mathbf{G}^T - \mathbf{d}^T\right)\cdot(\mathbf{G}\cdot\mathbf{m} - \mathbf{d}) = \mathbf{m}^T\mathbf{G}^T\mathbf{G}\mathbf{m} - \mathbf{m}^T\mathbf{G}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\mathbf{m} + \mathbf{d}^T\mathbf{d}, \tag{5.131}$$

where we have used the rule for the transpose of a matrix product (5.36). Note that since *E* is a scalar then each of these terms must evaluate to scalars as well. To find the solution, we set

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{G}\mathbf{m} + \mathbf{m}^T\mathbf{G}^T\mathbf{G}\dot{\mathbf{m}} - \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{d} - \mathbf{d}^T\mathbf{G}\dot{\mathbf{m}} = 0, j = 1, m, \tag{5.132}$$

where the "dot" over a vector represents the derivative of that vector with respect to m_j. We note the first and second terms are transposes of each other, as are the third and fourth terms. However, since they all evaluate to scalars the two transposes must be identical and hence this repetition simply constitutes a factor of two, which we delete by retaining only the first and third term:

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{G}\mathbf{m} - \dot{\mathbf{m}}^T\mathbf{G}^T\mathbf{d} = 0, j = 1, m. \tag{5.133}$$

What is the mysterious "dot"-derivative, written as

$$\dot{\mathbf{m}}^T = \frac{\partial}{\partial m_j}\left(\mathbf{m}^T\right), j = 1, m? \tag{5.134}$$

LINEAR (MATRIX) ALGEBRA

We illuminate this term by trying some values of *j*, remembering $\mathbf{m}^T = [m_1\ m_2 \cdots m_m]$:

$$\begin{aligned}
&\text{Case } j=1: \frac{\partial}{\partial m_1}\mathbf{m} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T \\
&\text{Case } j=2: \frac{\partial}{\partial m_2}\mathbf{m} = \begin{bmatrix} 0 & 1 & \cdots & 0 \end{bmatrix}^T \\
&\qquad\qquad \vdots \\
&\text{Case } j=m: \frac{\partial}{\partial m_m}\mathbf{m} = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}^T
\end{aligned}$$

Thus, the *m* linear equations can be combined into a single matrix equation, noting that all these derivatives (each producing a row vector) combine to form the identity matrix, I:

$$\frac{\partial}{\partial m_j}\left(\mathbf{m}^T\right), j=1,m \rightarrow \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}. \tag{5.135}$$

Hence, we may write

$$\mathbf{I}\mathbf{G}^T\mathbf{G}\mathbf{m} - \mathbf{I}\mathbf{G}^T\mathbf{d} = \mathbf{0}, \tag{5.136}$$

or by rearranging,

$$\mathbf{G}^T\mathbf{G}\mathbf{m} = \mathbf{G}^T\mathbf{d}. \tag{5.137}$$

Because the $\mathbf{G}^T\mathbf{G}$ matrix is square and symmetric and thus can be inverted, we simply multiply by its inverse and obtain the general least squares solution as

$$\mathbf{m} = \left[\mathbf{G}^T\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{d}. \tag{5.138}$$

LINEAR (MATRIX) ALGEBRA

Comparing (5.138) with (5.125) we see clearly that $N = G^T G$ and $v = G^T d$. Furthermore, given (5.128) we may write $G^T G$ using the product

$$\mathbf{N} = \mathbf{G}^T\mathbf{G} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_m^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_m \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1^T\mathbf{g}_1 & \mathbf{g}_1^T\mathbf{g}_2 & \cdots & \mathbf{g}_1^T\mathbf{g}_m \\ \mathbf{g}_2^T\mathbf{g}_1 & \mathbf{g}_2^T\mathbf{g}_2 & \cdots & \mathbf{g}_2^T\mathbf{g}_m \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{g}_m^T\mathbf{g}_1 & \mathbf{g}_m^T\mathbf{g}_2 & \cdots & \mathbf{g}_m^T\mathbf{g}_m \end{bmatrix} \tag{5.139}$$

Which makes it clear that each element of **N**, such as n_{kj}, is the dot product between two basis vectors $\mathbf{g}^T_k$ and $\mathbf{g}_j$, and

$$\mathbf{v} = \mathbf{G}^T\mathbf{d} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_m^T \end{bmatrix} \cdot \mathbf{d} = \begin{bmatrix} \mathbf{g}_1^T\mathbf{d} \\ \mathbf{g}_2^T\mathbf{d} \\ \vdots \\ \mathbf{g}_m^T\mathbf{d} \end{bmatrix}, \tag{5.140}$$

Which shows each element of **v** is the dot product between each basis function $\mathbf{g}^T_k$ and the data vector **d**. This is simply what we found the hard way earlier (i.e., 5.124). Thus, to solve a general linear least squares problem, all we have to do is to evaluate **G** via (5.128) and the rest is taken care of by (5.138).

LINEAR (MATRIX) ALGEBRA

Example 5–1. We are given a data set with four data pairs (2,1), (4,4), (6,3) and (8,4) (n = 4) and asked to determine the coefficients for a *quadratic* curve that best describes the data. Except for special situations, we know that the three-parameter curve will not pass through all the four points, so we decide to seek a least squares solution.

We write down the functional form for our quadratic curve as $d = m_1 + m_2 x + m_3 x^2$ and use it to form the matrix equation

$$\begin{aligned} m_1 + m_2 x_1 + m_3 x_1^2 &= d_1 \\ m_1 + m_2 x_2 + m_3 x_2^2 &= d_2 \\ m_1 + m_2 x_3 + m_3 x_3^2 &= d_3 \\ m_1 + m_2 x_4 + m_3 x_4^2 &= d_4 \end{aligned} \qquad (5.141)$$

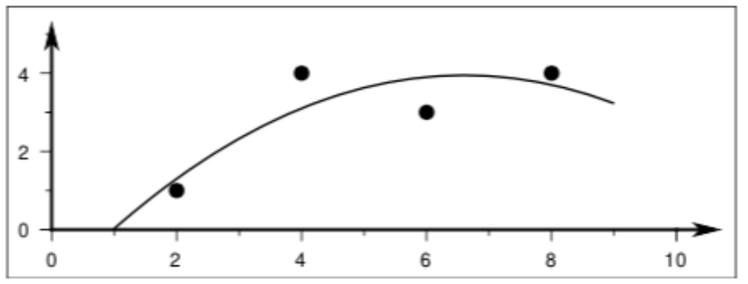

Figure 5.7: Fitting a three-parameter quadratic curve to four data points by minimizing the least squares misfit.

which for our data yields the linear system

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \\ 1 & 8 & 64 \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 4 \end{bmatrix}. \qquad (5.142)$$

Using (5.138) we find the solution to be $d = -1.5 + 1.65x - 0.125x^2$. Figure 5.7 shows our data as well as the best quadratic curve.

5.9.4 Weighted least squares solution

What if some data constraints are more reliable than others? We may simply give those residuals more weight than the others, i.e.,

$$\mathbf{e} = \begin{bmatrix} e_1 \\ 2e_2 \\ \vdots \\ e_n \end{bmatrix} \qquad (5.143)$$

In general, we can assign weights w_i to each misfit so that the new weighted misfits become $e'_i = e_i \cdot w_i$. Very often, the weights will simply be $s_{ii} = 1/s_i$, where s_i is the one-sigma uncertainty in the i'th measurement d_i.

LINEAR (MATRIX) ALGEBRA

We implement such weighting by introducing a diagonal weight matrix,

$$\mathbf{S} = \begin{bmatrix} s_{11} & & & \\ & s_{22} & & \\ & & \ddots & \\ & & & s_{nn} \end{bmatrix}, \tag{5.144}$$

which means the weighted residuals are **S · e** and the sum of the squared errors, *E* , becomes

$$E = (\mathbf{S}\cdot\mathbf{e})^T(\mathbf{S}\cdot\mathbf{e}) = \mathbf{e}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{e} = \mathbf{e}^T\cdot\mathbf{W}\cdot\mathbf{e}, \tag{5.145}$$

where we have introduced **W = S^{T}S**. Since **S·e = S·(G·m−d)** we obtain

$$\begin{aligned} E(\mathbf{m}) &= (\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d})^T\cdot(\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d}) = (\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T-\mathbf{d}^T\cdot S^{\mathbf{T}})\cdot(\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{S}\cdot\mathbf{d}) \\ &= \mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{d}^T\cdot\mathbf{S}^T\cdot\mathbf{S}\cdot\mathbf{d}. \end{aligned} \tag{5.146}$$

We substitute **W = S^{T}S**, take the partial derivatives, and obtain

$$\frac{\partial E(\mathbf{m})}{\partial m_j} = \mathbf{0} = \dot{\mathbf{m}}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\dot{\mathbf{m}}-\dot{\mathbf{m}}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\dot{\mathbf{m}},\ j=1,m. \tag{5.147}$$

Since **m** only contains the m_j, we know **$\dot{\mathbf{m}}^T$ = $\dot{\mathbf{m}}$ = I**. We again find the *m* normal equations can be written more compactly as the single matrix equation

$$\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}+\mathbf{m}^T\cdot\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}-\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}-\mathbf{d}^T\cdot\mathbf{W}\cdot\mathbf{G}=\mathbf{0}. \tag{5.148}$$

As before, the second and fourth terms are the transposes of the first and third terms, and as they all represent scalars our equation reduces to

$$\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\cdot\mathbf{m}-\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}=\mathbf{0}. \tag{5.149}$$

Thus, the weighted linear least squares solution is

$$\mathbf{m} = \left[\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{G}\right]^{-1}\mathbf{G}^T\cdot\mathbf{W}\cdot\mathbf{d}. \tag{5.150}$$

This solution is universal and applies to *any* linear least squares problem one can imagine. In the particular case when all s_{ii} = 1 the solution reduces to (5.138).

LINEAR (MATRIX) ALGEBRA

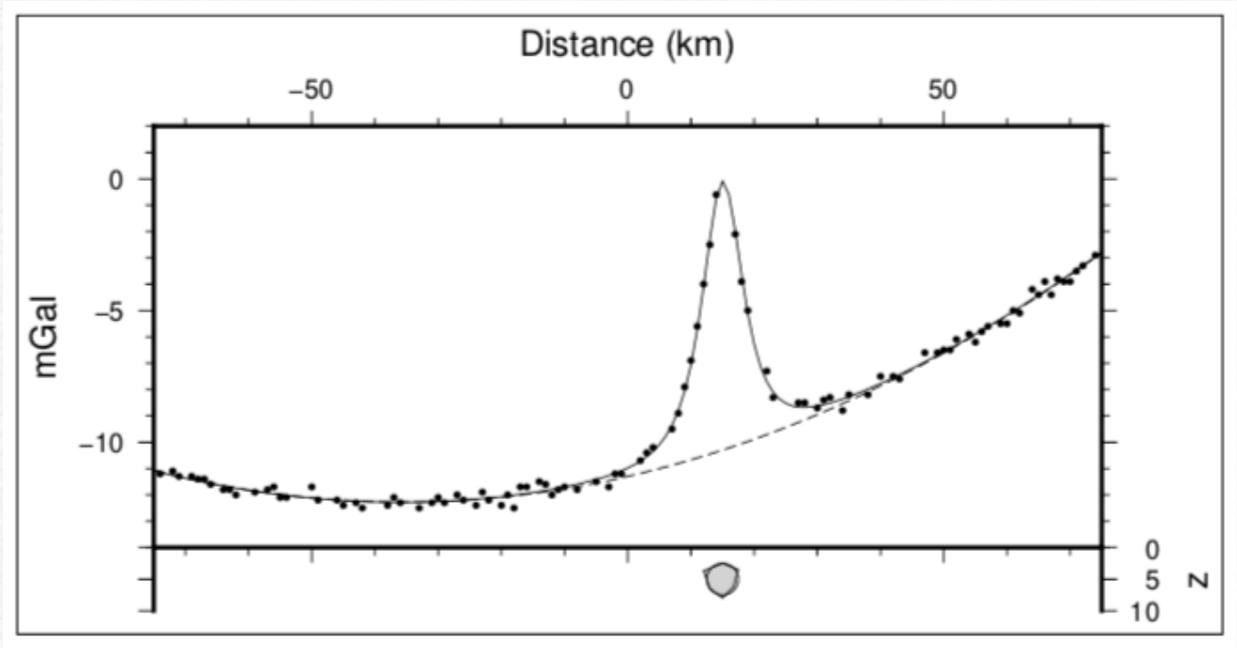

Figure 5.8: Observed and modeled gravity anomalies over a dense ore body. While the model is nonlinear in x and z, it is *linear* in the coefficients p_i.

Example 5–2. We will try the least squares machinery on an example taken from exploration geophysics. Figure 5.8 shows how observed gravity anomalies (d_i, solid circles) vary over a buried dense ore body as a function of location x_i. Based on the inferred geometry of the ore (from subsurface geology seen in mine shafts, etc.) we expect that a first-order approximation to the ore body could be a sphere buried at a depth of 5 km, with a radius of 2.5 km, and located at 15 km to the right of the origin. We would like to determine the density of that ore body. Exploration geophysics textbooks tell us that the gravity anomaly over a buried sphere of radius r and *unit* density is

$$g_{sp}(x,z,r) = \gamma \frac{\frac{4}{3}\pi r^3 z}{(x^2+z^2)^{3/2}}, \tag{5.151}$$

where z is the depth to the center of the sphere, $x = 0$ is where the sphere is located, and γ is the universal gravitational constant ($6.674 \cdot 10^{-11}\ m^3 kg^{-1} s^{-2}$). However, inspection of the data suggests that the anomaly due to the ore body is superimposed on a regional field with some curvature to it (i.e., dashed trend in Figure 5.8). Therefore, we decide to model these anomalies as a sum of a quadratic background (regional) field and the attraction of the sphere; this is accomplished with the four-parameter linear model

$$g(x) = m_1 + m_2 x + m_3 x^2 + m_4 g_{sp}(x-15,5,2.5), \tag{5.152}$$

where $m_4 = \Delta\rho$, the density contrast between the ore and the host rock. In the parlance of the previous sections, our basis functions $g_j(x)$ are $\{1, x, x^2, \text{and } g_{sp}(x)\}$.

LINEAR (MATRIX) ALGEBRA

To solve the problem we need to evaluate the matrix equation **G·m = d**, i.e.,

$$\begin{bmatrix} 1 & x_1 & x_1^2 & g_{sp}(x_1-15,5,2.5) \\ 1 & x_2 & x_2^2 & g_{sp}(x_2-15,5,2.5) \\ 1 & x_3 & x_3^2 & g_{sp}(x_3-15,5,2.5) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & g_{sp}(x_n-15,5,2.5) \end{bmatrix} \cdot \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix} \tag{5.153}$$

whose solution becomes

$$\mathbf{m} = [\mathbf{G}^T \cdot \mathbf{G}]^{-1}\mathbf{G}^T \cdot \mathbf{d}. \tag{5.154}$$

Thus, we have solved a fairly complicated least squares modeling problem (solution is the solid line in Figure 5.8).