Excursions in linear least squares

Daniel Topa

E-mail address: dantopa@gmail.com

2010 *Mathematics Subject Classification.* Primary
Key words and phrases. least squares, L^2, l^2

Abstract.

Contents

Part 1

Rudiments

Part 2

Zonal fits

Part 3

Modal fits

Part 4

Nonlinear problems

CHAPTER 1

Strays

Blindly applying linear tools to nonlinear problems presents many paths to perdition. Hope, no matter how fervent, cannot remedy mathematical maladies.

We stress the definition of the least squares problem in (**??**) as the first indication that something is amiss. We stress visualization methods to help reveal the status of a calculation.

(1) finding reasonable approximations which nudge the problem into linearity;
(2) iterating the solution to a linear problem to improve a nonlinear problem;
(3) separating a problem into linear and nonlinear components.

1.1. Finding the best circle

1.1.1. Nonlinear formulation. This is an example of a numbered first-level heading.

A circle is characterized by two parameters: an origin and a radius. The origin is a vector quantity, the radius a scalar.

$$O = \left[\begin{array}{c} x_0 \\ y_0 \end{array} \right] \tag{1.1}$$

Given a set of measurements p_j, $j = 1\colon m$.

$$(x - x_0)^2 - (y - y_0)^2 = \rho^2 \tag{1.2}$$

This implies a trial function

$$\chi^2 (O, \rho) = \sum_{j=1}^{m} \left(\rho^2 - (x_j - x_0)^2 + (y_j - y_0)^2 \right)^2 \tag{1.3}$$

In equation (1.2) the fit parameters for the origin appear in a nonlinear fashion, making this a nonlinear problem. There are many ways to solve such a problem. However, our focus is on linear problems.

1.1.2. Linear formulation. We start with the simple vector equation

$$p_j = r_k + O \tag{1.4}$$

from which we conclude

$$p_j^2 = r_j^2 + O^2 + 2 r_j \cdot O \tag{1.5}$$

The trick is make one parameter disappear. To do so examine differences between the measurements

$$\Delta_{jk} = p_j - p_k = r_j - r_k \tag{1.6}$$

The data is no longer a list of m measurements of p; instead it is a list of τ differences where

$$\tau = \frac{1}{2}m(m-1) \tag{1.7}$$

For example, when $m = 4$

TABLE 1.1. The new data set compared to the old. The measured values p are converted to a set of differences Δ_{jk}.

	measurements	inputs
1	p_1	$\Delta_{12} = p_1 - p_2$
2	p_2	$\Delta_{13} = p_1 - p_3$
3	p_3	$\Delta_{14} = p_1 - p_4$
4	p_4	$\Delta_{23} = p_2 - p_3$
5		$\Delta_{24} = p_2 - p_4$
6		$\Delta_{34} = p_3 - p_4$

$$p_j^2 - p_k^2 = r_j^2 - r_k^2 + 2\left(r_j - r_k\right) \cdot O \tag{1.8}$$

$$r_j^2 = \rho^2 \qquad j = 1\colon m \tag{1.9}$$

$$r_j^2 - r_k^2 = 0 \qquad j,k = 1\colon m \tag{1.10}$$

The final trial function is this using equation (1.6)

$$p_j^2 - p_k^2 = 2\left(p_j - p_k\right) \cdot O \tag{1.11}$$

The trial function is then

$$p_j^2 - p_k^2 = 2\left(p_j - p_k\right) \cdot O \tag{1.12}$$

and the merit function

$$\chi^2\left(O\right) = \sum_{j=1}^{m-1}\sum_{k=1}^{m}\left(p_j^2 - p_k^2 - 2\left(p_j - p_k\right) \cdot O\right)^2 \tag{1.13}$$

Label the pairs

$$\xi = \left\{ \left[\begin{array}{c} 1 \\ 2 \end{array}\right], \left[\begin{array}{c} 1 \\ 3 \end{array}\right], \dots, \left[\begin{array}{c} m-1 \\ m \end{array}\right] \right\} \tag{1.14}$$

$$\chi^2\left(O\right) = \sum_{\mu=1}^{\tau}\left(2\Delta_\xi \cdot O - p_{\xi_1}^2 + p_{\xi_2}^2\right)^2 \tag{1.15}$$

Linear system

$$\begin{aligned} p_1^2 - p_2^2 &= 2\left(p_1 - p_2\right) \cdot O \\ p_1^2 - p_3^2 &= 2\left(p_1 - p_3\right) \cdot O \\ &\vdots \\ p_{m-1}^2 - p_m^2 &= 2\left(p_{m-1} - p_m\right) \cdot O \end{aligned} \tag{1.16}$$

k	$.x_k$	$.r_k$	$.x_k r_k$	$.g(x_k, r_k)$	.residual
1	0.0600	15.78246300	0.94694800	0.87847100	0.068476600
2	0.0500	18.89113500	0.94455700	0.88661400	0.057942900
3	0.0400	23.96075200	0.9584300	0.89475600	0.063673600
4	0.0300	32.35313500	0.97059400	0.90289900	0.067694900
5	0.0200	50.5914400	1.011828800	0.91104200	0.10078700
6	0.0100	104.68871700	1.046887200	0.91918400	0.12770300
7	0.0000	−1839.04936400	0.0000	0.92732700	−0.92732700
8	−0.0100	−103.18431800	1.031843200	0.9354700	0.096373500
9	−0.0200	−50.73661200	1.014732200	0.94361200	0.071119900
10	−0.0300	−33.75893700	1.012768100	0.95175500	0.061013100
11	−0.0400	−25.71153700	1.028461500	0.95989800	0.068563800
12	−0.0500	−20.80382100	1.040191100	0.9680400	0.072150700
13	−0.0600	−17.46685300	1.048011200	0.97618300	0.071828200

TABLE 1.2. Raw data for focal length measurement

solve for the origin O. The problem statement

$$\Delta O = b \tag{1.17}$$

In d dimensions the matrix dimensions are

$$\Delta \in \mathbb{R}_d^{\tau\times d}, \quad O \in \mathbb{R}^{d\times 1}, \quad b \in \mathbb{R}^{\tau\times 1}$$

and the matrices are defined as

$$\Delta = 2\begin{bmatrix} p_1 - p_2 \\ \vdots \\ p_{m-1} - p_m \end{bmatrix}, \quad O = \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}, \quad b = \begin{bmatrix} p_1^2 - p_2^2 \\ \vdots \\ p_{m-1}^2 - p_m^2 \end{bmatrix} \tag{1.18}$$

1.2. Dubious applications

Let's explore the application of linear methods to nonlinear problems. Laboratory constrains mathematics. You inherit a spreadsheet and are asked to do basic analysis.

1.2.1. First analysis. Thin lens equation

$$x\left(r + e\right) = -f^2 \tag{1.19}$$

Trial function

$$\phi + ex = -xr \tag{1.20}$$

Physical fact

$$|\phi| = f^2 \tag{1.21}$$

Linear system

$$\mathbf{A}z = b$$

$$\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \phi \\ e \end{bmatrix} = -\begin{bmatrix} x_1 R_1 \\ \vdots \\ x_m R_m \end{bmatrix} \tag{1.22}$$

measurements	$x_k,\ k = 1 : m$ $R_k,\ k = 1 : m$
fit parameters	$\phi \pm \epsilon_\phi$ $e \pm \epsilon_e$
trial function	$\phi + ex = -xR$
merit function	$\sum_{k=1}^{m} (\phi + ex_k + x_k R_k)^2$
linear system	$\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \phi \\ e \end{bmatrix} = - \begin{bmatrix} x_1 R_1 \\ \vdots \\ x_m R_m \end{bmatrix}$

TABLE 1.3. Problem statement: determine focal length f

fit parameters	$\phi \pm \epsilon_\phi = -0.927 \pm 0.081$ $e \pm \epsilon_e = 0.8 \pm 2.2$
$r^{\mathrm{T}} r$	0.08511
c	$\frac{1}{91} \begin{bmatrix} 7 & 0 \\ 0 & 5000 \end{bmatrix}$
plots	data vs fit (**??**) residuals (**??**) merit function in $\mathcal{R}(\mathbf{A}^*)$ (**??**)

TABLE 1.4. Results: focal length f

The expected focal length is $f = 1$ m.

$$f_{measured} = 0.963 \pm 0.062 \text{ m} \tag{1.23}$$

k	$.x_k$	$.r_k$	$.x_k r_k$	$.g(x_k, r_k)$	.residual
1	0.0600	15.78246300	0.94694800	−0.95574800	−0.0088006200
2	0.0500	18.89113500	0.94455700	−0.96389100	−0.019334300
3	0.0400	23.96075200	0.9584300	−0.97203400	−0.013603600
4	0.0300	32.35313500	0.97059400	−0.98017600	−0.0095823200
5	0.0200	50.5914400	1.011828800	−0.98831900	0.023509800
6	0.0100	104.68871700	1.046887200	−0.99646200	0.050425500
7	−0.0100	−1839.04936400	18.39049400	−1.01274700	0.019096200
8	−0.0200	−103.18431800	2.063686400	−1.020889600	−0.0061573900
9	−0.0300	−50.73661200	1.52209800	−1.029032300	−0.016264200
10	−0.0400	−33.75893700	1.35035700	−1.037174900	−0.0087134400
11	−0.0500	−25.71153700	1.28557700	−1.045317600	−0.0051265100
12	−0.0600	−20.80382100	1.24822900	−1.053460200	−0.0054490700

TABLE 1.5. Input data for focal length measurement

fit parameters	$\phi \pm \epsilon_\phi = -1.0046 \pm 0.0062$
	$e \pm \epsilon_e = 0.81 \pm 0.16$
$r^{\mathrm{T}} r$	0.004623
c	$\frac{1}{1092}\left[\begin{array}{cc} 91 & 0 \\ 0 & 60\,000 \end{array}\right]$
plots	data vs fit (**??**)
	residuals (**??**)
	merit function in $\mathcal{R}(\mathbf{A}^*)$ (1.3)

TABLE 1.6. Improved results: focal length f

1.2.2. Second analysis.

$$f_{measured} = 1.0023 \pm 0.0042 \text{ m} \tag{1.24}$$

Precision improves by an order of magnitude when the point at the origin is excluded. The exclusion criteria is based on the statistics of the data set, not difficulty in the measurement.

1.3. Population growth

In this section we take a nonlinear model for population growth and separate the linear and nonlinear terms.

1.3.1. Population growth.

$$y(t) = c_1 + c_2 (t - 1900) + c_3 e^{d(t-1900)} \tag{1.25}$$

$$\min_{c \in \mathbb{R}^3} \left\| \mathbf{A}(d) \left[\begin{array}{c} c_1 \\ c_2 \\ c_3 \end{array} \right] - y \right\|_2^2 \tag{1.26}$$

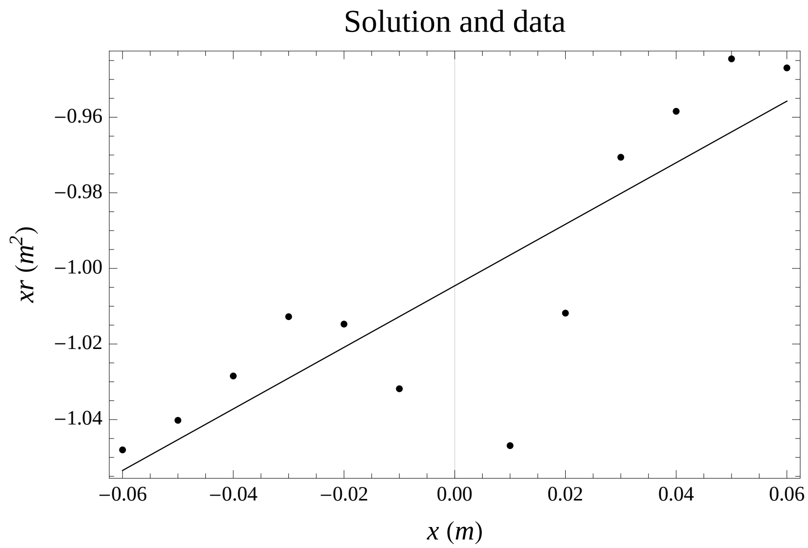

FIGURE 1.1. The solution curve (**??**) plotted against the data in table (**??**).

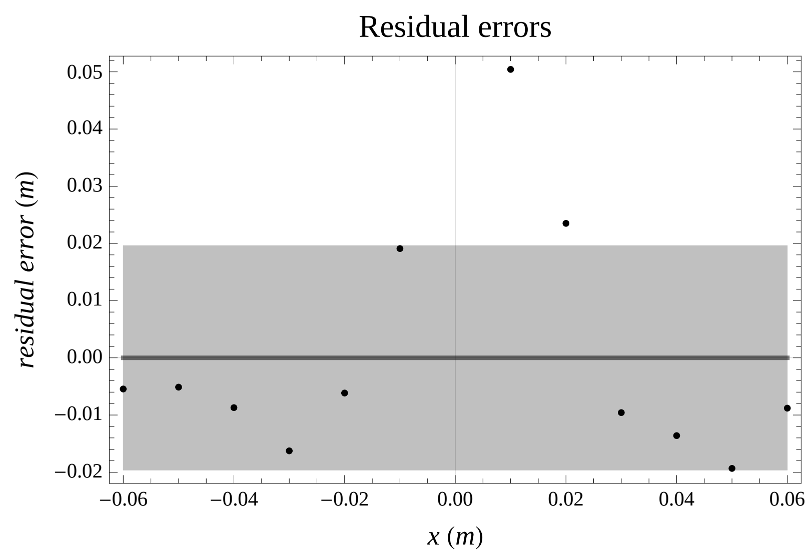

FIGURE 1.2. A closer look at the residual errors plotted on an absolute scale.

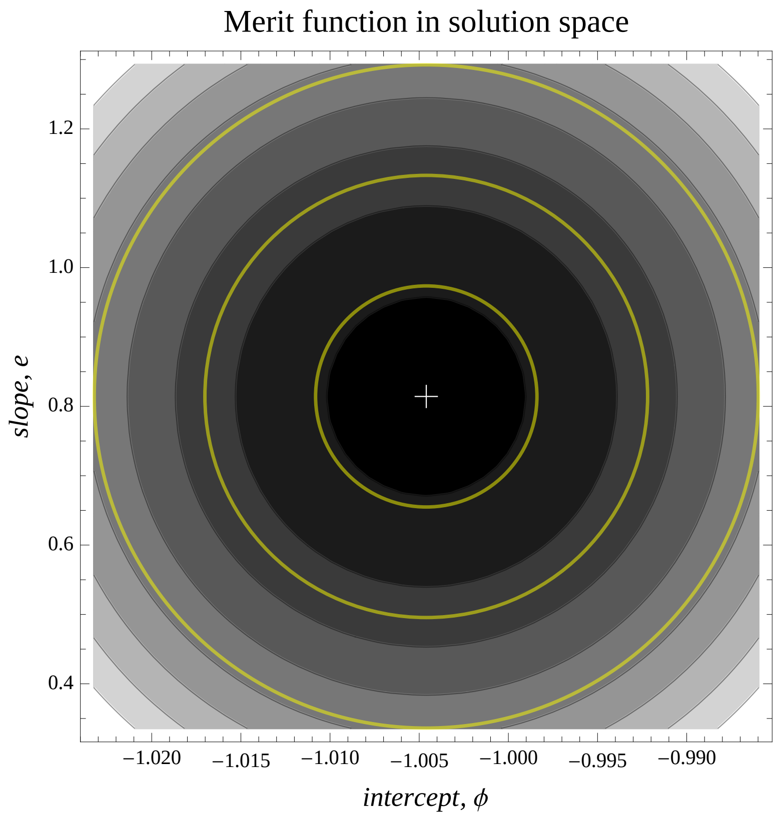

FIGURE 1.3. The merit function in solution space: the white cross in the center marks the solution found in (**??**), the yellow curves represent the error ellipses with radii of (a_0, a_1), $2\,(a_0, a_1)$, and $3\,(a_0, a_1)$.

year	census	fit	r	rel err
1900	76.00	77.51	1.51	2.0%
1910	91.97	90.98	-0.99	-1.1%
1920	105.71	104.87	-0.84	-0.8%
1930	122.78	119.48	-3.29	-2.7%
1940	131.67	135.36	3.69	2.8%
1950	150.70	153.46	2.76	1.8%
1960	179.32	175.45	-3.87	-2.2%
1970	203.24	204.26	1.029	0.5%

TABLE 1.7. Data v. prediction.

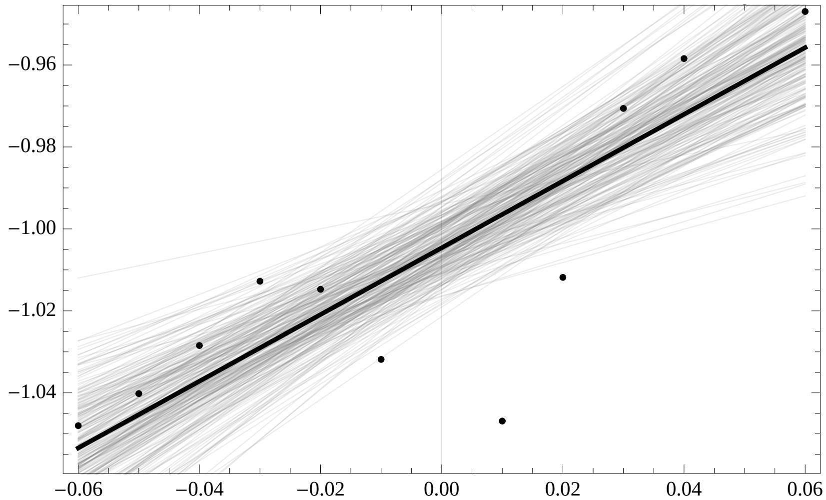

FIGURE 1.4. Whisker plot.

fit parameters	$c = \begin{bmatrix} 0.010 \\ 0.0170 \\ 0.0096 \end{bmatrix} \pm \begin{bmatrix} 0.031 \\ 0.0014 \\ 0.0020 \end{bmatrix}$
	$d = 0.056136 \pm ?.?$
$r^{\mathrm{T}} r$	0.009025
α	$\begin{bmatrix} 0.5397 & -0.0188 & 0.0165 \\ -0.0188 & 0.0011 & -0.0014 \\ 0.0165 & -0.0014 & 0.0022 \end{bmatrix}$
plots	data vs fit (**??**) residuals (**??**) merit function in $\mathcal{R}(\mathbf{A}^*)$ (**??**)

TABLE 1.8. Results: census

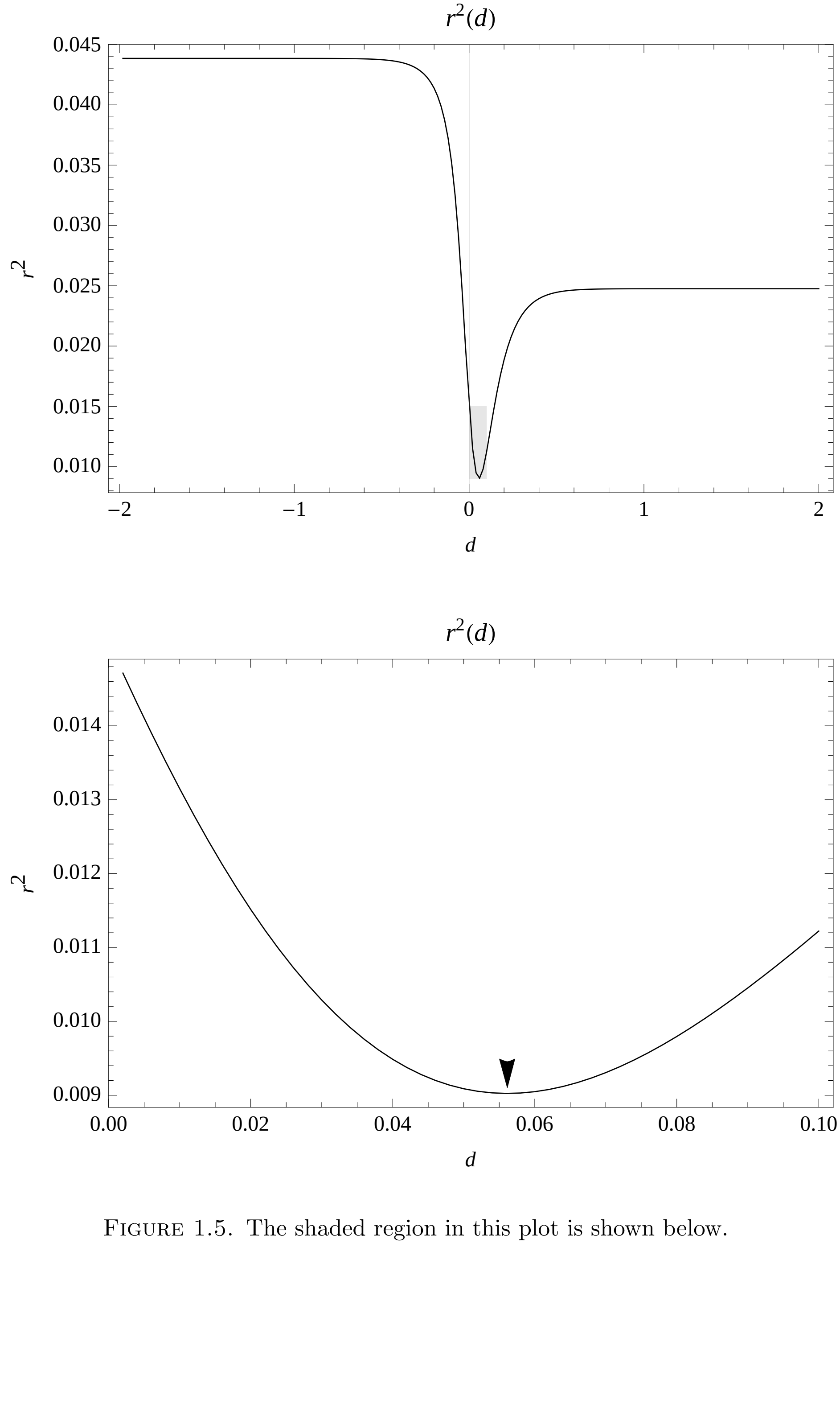

FIGURE 1.5. The shaded region in this plot is shown below.

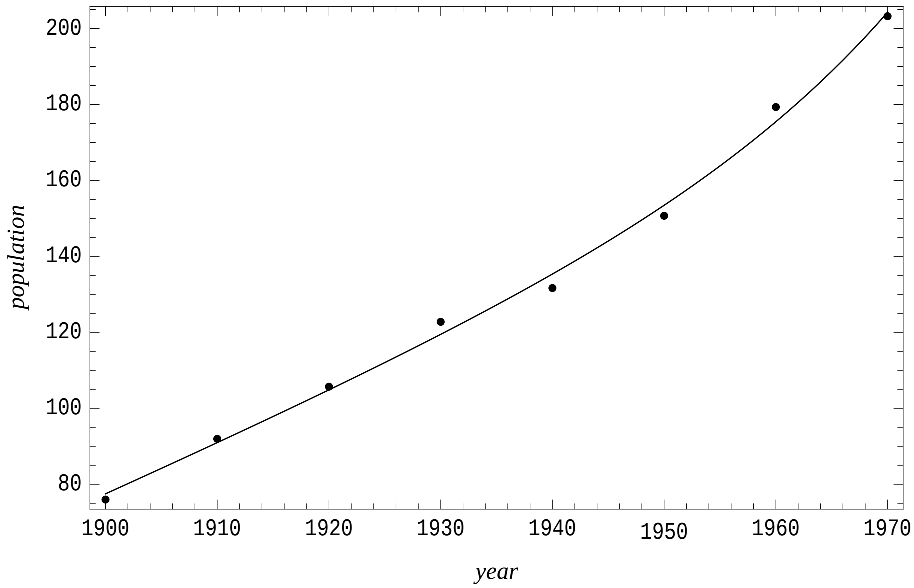

FIGURE 1.6. Solution plotted against data.

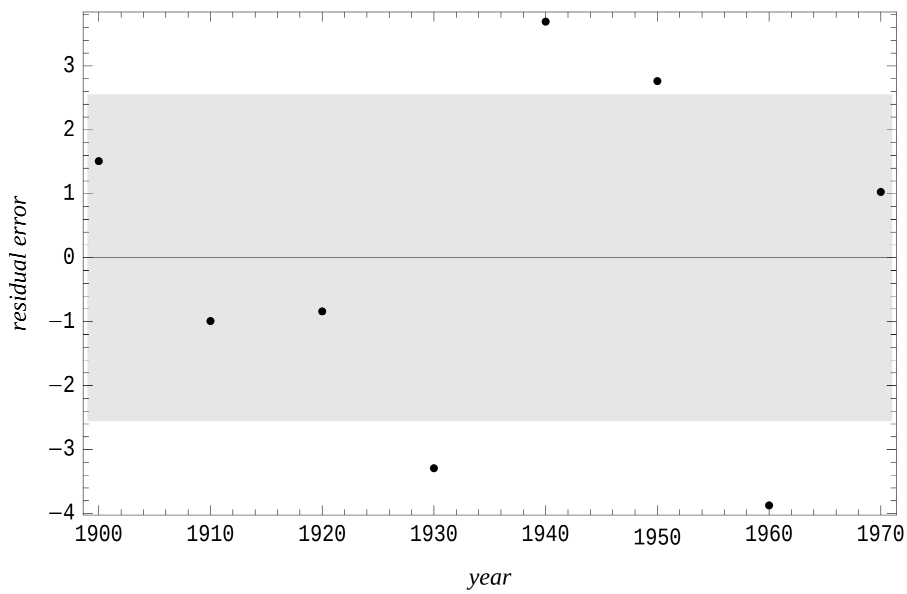

FIGURE 1.7. Residual errors.

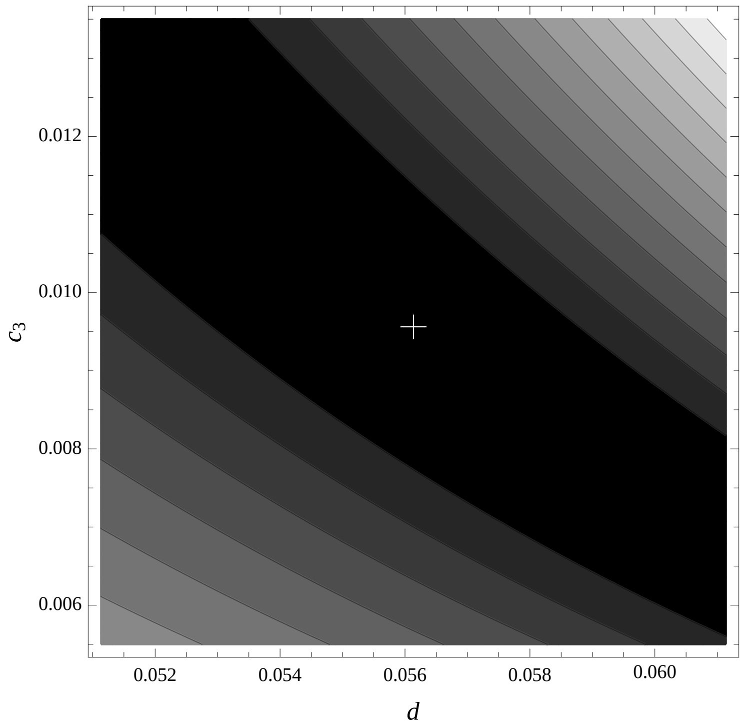

FIGURE 1.8. The merit function with c fixed at best values.

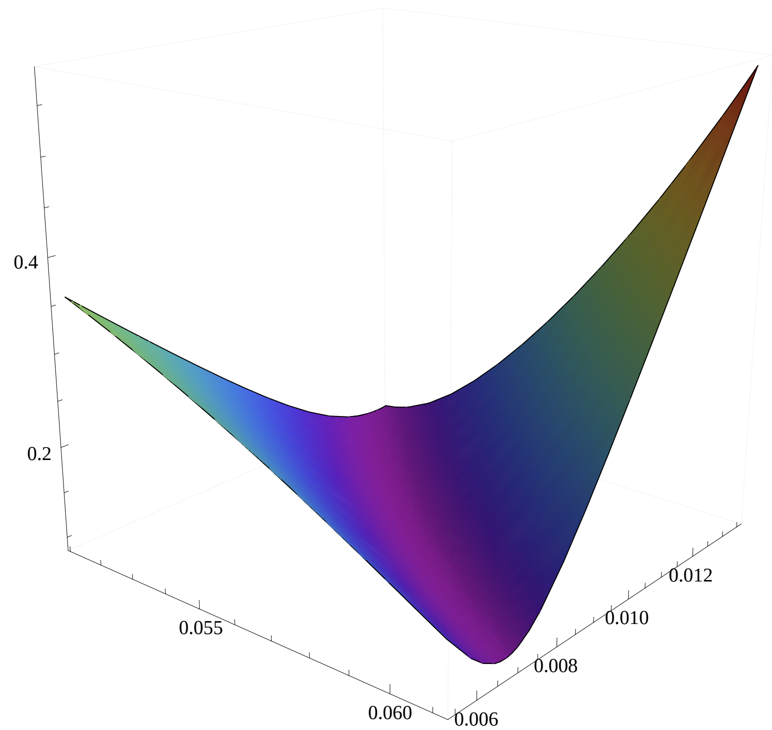

Figure 1.9. The merit function in three dimensions.

APPENDIX A

Least squares with exemplar matrices

A.1. Linear systems

The essential concepts of least squares and the fundamental subspaces spring to life using exemplar matrices. The canonical linear system is

$$\mathbf{A}x = b$$

The matrix $\mathbf{A}$ has m rows and n columns of complex numbers. The matrix rank is $\rho \le \min(m, n)$. In shorthand, the three components are

(1) $\mathbf{A} \in \mathbb{C}^{m \times n}_{\rho}$: the system matrix, an input;
(2) $b \in \mathbb{C}^{m}$: the data vector, an input;
(3) $x \in \mathbb{C}^{n}$: the solution vector, the output.

The residual error from the best fit is

$$r = \mathbf{A}x - b. \tag{A.1}$$

Ignore the trivial cases where $b = 0$.

Exemplar matrices have obvious singular value decompositions.

TABLE A.1. Exemplar matrices and their block forms.

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \quad \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]$$

$$\left[\begin{array}{c} \mathbf{I}_2 \end{array}\right] \quad \left[\begin{array}{c} \mathbf{I}_2 \\ \hline \mathbf{0} \end{array}\right] \quad \left[\begin{array}{c|c} \mathbf{I}_2 & \mathbf{0} \end{array}\right] \quad \left[\begin{array}{c|c} \mathbf{I}_2 & \mathbf{0} \\ \hline \mathbf{0} & 0 \end{array}\right]$$

A.2. Exemplars

A.2.1. Full rank: $\rho = m = n$. Start with an ideal linear system

$$\mathbf{A}x = b$$

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right]. \tag{A.2}$$

Subspace decomposition:

Because the matrix $\mathbf{A}$ has full column rank the null space $\mathcal{N}(\mathbf{A}^*)$ is trivial. Because the matrix $\mathbf{A}^*$ has full row rank the null space $\mathcal{N}(\mathbf{A})$ is trivial.

TABLE A.2. Exemplar matrices and their block forms.

exemplar	block form
$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$	$\left[\begin{array}{c} \mathbf{I}_2 \end{array}\right]$
$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right]$	$\left[\begin{array}{c} \mathbf{I}_2 \\ \hline \mathbf{0} \end{array}\right]$
$\left[\begin{array}{cc\|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$	$\left[\begin{array}{c\|c} \mathbf{I}_2 & \mathbf{0} \end{array}\right]$
$\left[\begin{array}{cc\|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right]$	$\left[\begin{array}{c\|c} \mathbf{I}_2 & \mathbf{0} \\ \hline \mathbf{0} & 0 \end{array}\right]$

$$\begin{array}{lcccccccc} \text{domain:} & \mathbb{C}^2 & = & \mathcal{R}(\mathbf{A}^*) & \oplus & \mathcal{N}(\mathbf{A}) & = & \text{sp}\left\{\left[\begin{array}{c}1\\0\end{array}\right],\left[\begin{array}{c}0\\1\end{array}\right]\right\} & \oplus & \text{sp}\left\{\left[\begin{array}{c}0\\0\end{array}\right]\right\} \\ \text{codomain:} & \mathbb{C}^2 & = & \mathcal{R}(\mathbf{A}) & \oplus & \mathcal{N}(\mathbf{A}^*) & = & \text{sp}\left\{\left[\begin{array}{c}1\\0\end{array}\right],\left[\begin{array}{c}0\\1\end{array}\right]\right\} & \oplus & \text{sp}\left\{\left[\begin{array}{c}0\\0\end{array}\right]\right\} \end{array}$$

TABLE A.3. Subspace decomposition for the **A** matrix in (A.2).

Existence and uniqueness: We have unconditional existence and uniqueness without regard to the data vector. The exact solution is

$$x = \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.3}$$

which is also the least squares solution

$$x_{LS} = x = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.4}$$

with $r^{\mathrm{T}}r = 0$ residual error. More formally, the linear system has a unique solution for any value of $b_1, b_2 \in \mathbb{C}$.

A.2.2. Full column rank: $\rho = n < m$. Foreshadowing the resolution of the range and null spaces, we show a partitioning

$$\begin{array}{c} \mathbf{A}x = b \\ \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \\ \hline b_3 \end{array}\right]. \end{array} \tag{A.5}$$

Subspace decomposition:

Fundamental Theorem of Linear Algebra

$$\begin{array}{llllllll} \text{domain:} & \mathbb{C}^2 & = & \mathcal{R}(\mathbf{A}^*) & \oplus & \mathcal{N}(\mathbf{A}) & = & \text{sp}\left\{ \left[\begin{array}{c}1\\0\end{array}\right], \left[\begin{array}{c}0\\1\end{array}\right] \right\} \oplus \text{sp}\left\{ \left[\begin{array}{c}0\\0\end{array}\right] \right\} \\ \text{codomain:} & \mathbb{C}^3 & = & \mathcal{R}(\mathbf{A}) & \oplus & \mathcal{N}(\mathbf{A}^*) & = & \text{sp}\left\{ \left[\begin{array}{c}1\\0\\0\end{array}\right], \left[\begin{array}{c}0\\1\\0\end{array}\right] \right\} \oplus \text{sp}\left\{ \left[\begin{array}{c}0\\0\\1\end{array}\right] \right\} \end{array}$$

statement	subspace condition	data conditions
existence and uniqueness	$b \in \mathcal{R}(\mathbf{A})$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 = 0$
existence	$b \in \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^*)$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 \neq 0$
no existence	$b \in \mathcal{N}(\mathbf{A}^*)$	$b_1 = b_2 = 0$, $b_3 \in \mathbb{C}$

TABLE A.4. Existence and uniqueness for the full column rank linear system in equation (A.5).

Thanks to the gentle behavior of the exemplar matrix, the range and null space components for the data vector are apparent:

$$b = \underbrace{\left[\begin{array}{c} b_1 \\ b_2 \\ 0 \end{array}\right]}_{\in \mathcal{R}(\mathbf{A})} + \underbrace{\left[\begin{array}{c} 0 \\ 0 \\ b_3 \end{array}\right]}_{\in \mathcal{N}(\mathbf{A}^*)} \tag{A.6}$$

Existence and uniqueness: When the data vector component $b_3 = 0$,

$$b = \left[\begin{array}{c} b_1 \\ b_2 \\ 0 \end{array}\right] \in \mathcal{R}(\mathbf{A}) \tag{A.7}$$

the linear system is consistent and we have a unique solution

$$x = \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.8}$$

which is also the least squares solution

$$x_{LS} = x = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.9}$$

with $r^{\mathrm{T}} r = 0$ residual error. Notice that the solution vector is in the complementary range space, the range space of $\mathbf{A}^*$:

$$x \in \mathcal{R}(\mathbf{A}^*). \tag{A.10}$$

No existence: When the data vector inhabits the null space we do not even have a least squares solution.

Existence, no uniqueness:

A.2.3. Full row rank: $\rho = m < n$**.** Foreshadowing the resolution of the range and null spaces, we show a partitioning

$$\mathbf{A}x = b$$

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right]. \tag{A.11}$$

Subspace decomposition:

Fundamental Theorem of Linear Algebra

$$\begin{array}{lcccccccc} \text{domain:} & \mathbb{C}^3 & = & \mathcal{R}(\mathbf{A}^*) & \oplus & \mathcal{N}(\mathbf{A}) & = & \text{sp}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} & \oplus \ \text{sp}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \\ \text{codomain:} & \mathbb{C}^2 & = & \mathcal{R}(\mathbf{A}) & \oplus & \mathcal{N}(\mathbf{A}^*) & = & \text{sp}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} & \oplus \ \text{sp}\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \end{array}$$

Thanks to the gentle behavior of the exemplar matrix, the range and null space components for the solution vector are apparent:

$$x = \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}}_{\in \mathcal{R}(\mathbf{A}^*)} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix}}_{\in \mathcal{N}(\mathbf{A})} \tag{A.12}$$

Existence and uniqueness: When the data vector component $b_3 = 0$,

$$b = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \in \mathcal{R}(\mathbf{A}) \tag{A.13}$$

the linear system is consistent and we have a unique solution

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{A.14}$$

which is also the least squares solution

$$x_{LS} = x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{A.15}$$

with $r^{\mathrm{T}} r = 0$ residual error. Notice that the solution vector is in the complementary range space, the range space of $\mathbf{A}^*$:

$$x \in \mathcal{R}(\mathbf{A}^*). \tag{A.16}$$

No existence When the data vector inhabits the null space we do not even have a least squares solution.

Existence, no uniqueness:

A.2.4. Row and column rank deficit: $\rho < m, \rho < n$. Partitioning

$$\mathbf{A}x = b$$

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \tag{A.17}$$

Singular Value Decomposition

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] = \mathbf{U}\,\Sigma\,\mathbf{V}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{A.18}$$

Subspace decomposition:

statement	subspace condition	data conditions
existence and uniqueness	$b \in \mathcal{R}(\mathbf{A})$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 = 0$
existence	$b \in \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^*)$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 \neq 0$
no existence	$b \in \mathcal{N}(\mathbf{A}^*)$	$b_1 = b_2 = 0$, $b_3 \in \mathbb{C}$

TABLE A.5. Existence and uniqueness for the full column rank linear system in equation (A.5).

Fundamental Theorem of Linear Algebra

$$\text{domain:} \quad \mathbb{C}^3 = \mathcal{R}(\mathbf{A}^*) \oplus \mathcal{N}(\mathbf{A}) = \text{sp}\left\{ \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right], \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] \right\} \oplus \text{sp}\left\{ \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] \right\}$$

$$\text{codomain:} \quad \mathbb{C}^3 = \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^*) = \text{sp}\left\{ \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right], \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] \right\} \oplus \text{sp}\left\{ \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] \right\}$$

Thanks to the gentle behavior of the exemplar matrix, the range and null space components for the solution vector are apparent:

$$x = \underbrace{\left[\begin{array}{c} x_1 \\ x_2 \\ 0 \end{array}\right]}_{\in \mathcal{R}(\mathbf{A}^*)} + \underbrace{\left[\begin{array}{c} 0 \\ 0 \\ x_3 \end{array}\right]}_{\in \mathcal{N}(\mathbf{A})} \tag{A.19}$$

Existence and uniqueness: When the data vector component $b_3 = 0$,

$$b = \left[\begin{array}{c} b_1 \\ b_2 \\ 0 \end{array}\right] \in \mathcal{R}(\mathbf{A}) \tag{A.20}$$

the linear system is consistent and we have a unique solution

$$x = \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.21}$$

which is also the least squares solution

$$x_{LS} = x = \left[\begin{array}{c} b_1 \\ b_2 \end{array}\right] \tag{A.22}$$

with $r^{\mathrm{T}} r = 0$ residual error. Notice that the solution vector is in the complementary range space, the range space of $\mathbf{A}^*$:

$$x \in \mathcal{R}(\mathbf{A}^*). \tag{A.23}$$

No existence When the data vector inhabits the null space we do not even have a least squares solution.

Existence, no uniqueness:

statement	subspace condition	data conditions
existence and uniqueness	$b \in \mathcal{R}(\mathbf{A})$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 = 0$
existence	$b \in \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^*)$	$(b_1 \neq 0$ or $b_2 \neq 0)$ and $b_3 \neq 0$
no existence	$b \in \mathcal{N}(\mathbf{A}^*)$	$b_1 = b_2 = 0$, $b_3 \in \mathbb{C}$

TABLE A.6. Existence and uniqueness for the full column rank linear system in equation (A.5).