WORCESTER POLYTECHNIC INSTITUTE MECHANICAL ENGINEERING DEPARTMENT

Engineering Experimentation ME-3901, D'2012

Lecture 04 21 March 2012





General information

Office hours

<u>Instructors</u>: Cosme Furlong Office: HL-151 <u>Everyday</u>: 9:00 to 9:50 am Christopher Scarpino Office: HL-153 During laboratory sessions

<u>Teaching Assistants</u>: During laboratory sessions





Curve fitting: introduction

The most common and most efficient ways to present quantitative data are by the use of tables, plots (figures), and mathematical equations. Which method of presentation is the most appropriate form of reporting depends on a variety of factors, which include the purpose of the presentation, the quantity of data, the degree of precision of the results, and the number of variables. Because so many factors are involved, only a few general remarks can be made. Your choice of the particular form of presentation must be made largely by applying common sense guided by experience.

Table 6.1.			Table 6.2.			Tabl	e 6.3.		Table 6.4.	
Х	Y		Х	Y		Х	Y		Х	Y
10.0	8.04		10.0	9.14		10.0	7.46		8.0	6.58
8.0	6.95		8.0	8.14		8.0	6.77		8.0	5.76
13.0	7.58		13.0	8.74		13.0	12.74		8.0	7.71
9.0	8.81		9.0	8.77		9.0	7.11		8.0	8.84
11.0	8.33		11.0	9.26		11.0	7.81		8.0	8.47
14.0	9.96		14.0	8.10		14.0	8.84	-	8.0	7.04
6.0	7.24		6.0	6.13		6.0	6.08		8.0	5.25
4.0	4.26		4.0	3.10		4.0	5.39		19.0	12.50
12.0	10.84		12.0	9.13		12.0	8.15		8.0	5.56
7.0	4.82		7.0	7.26		7.0	6.42		8.0	7.91
5.0	5.68		5.0	4.74		5.0	5.73		8.0	6.89

For example, four sets of data are shown in Tables 6.1 to 6.4, respectively. Each one of these data sets has exactly the same mean value of X's = 9.0 and of Y's = 7.5. Their standard deviations are also exactly the same and equal to $\sigma_x = 3.16$ and $\sigma_y = 1.94$. Furthermore, each of these four data sets is represented by the same equation of a straight line, obtained by the least-squares analysis, which is

$$Y = 0.50X + 3.00 \quad , \tag{6.1}$$

and which results in a correlation coefficient of 0.77, for each data set. However, graphical representations of the four data sets are drastically different, as shown in Figs







The foregoing discussion indicates that mere tabulation of the data and computation of statistical parameters representing these data are not adequate for meaningful presentation of results. Graphical representation of the results is extremely important and aided with statistics and mathematical equations provides complete presentation of data. The subject of this note is graphical, statistical, and mathematical presentation of data and is illustrated with a representative problem.

6.2. Method of least squares

Table 6.1 represents variables X and Y over a range of values. What is needed is a mathematical expression that will represent Y as a function of X. The simplest type of function is a linear one, which establishes Y as a linear function of X expressed as



 $Y=aX+b \quad ,$







Simplifications	on of Eqs. 6.5 and 6.6 results in a system of two simultaneous	ous
	$2\sum_{i=1}^{n} [(aX_i + b)X_i] = 2\sum_{i=1}^{n} (Y_iX_i),$	(6.3
and		
	$2\sum_{i=1}^{n} (aX_i + b) = 2\sum_{i=1}^{n} Y_i ,$	(6.8
which can be rewrit	iten as	
	$a \sum_{i=1}^{n} (X_i)^2 + b \sum_{i=1}^{n} (X_i) = \sum_{i=1}^{n} (Y_i X_i),$	(6.9
and	$a\sum_{i=1}^{n} (\mathbf{X}_i) + bn = \sum_{i=1}^{n} (\mathbf{Y}_i).$	(6.10
Equations 6.	9 and 6.10 can be written in a matrix form as	
	$\begin{bmatrix} \sum_{i=1}^{n} X_i^2 & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & n \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} X_i Y_i \\ \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} Y_i \end{pmatrix} .$	(6.11
Solution of Eq. 6.11	yields	
	$ \begin{pmatrix} a \\ b \end{pmatrix} = \begin{bmatrix} \sum_{i=1}^{n} X_i^2 & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & n \end{bmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n} X_i Y_i \\ \sum_{i=1}^{n} Y_i \end{pmatrix} . $	(6.12
or		
$\left(\begin{array}{c}a\\b\end{array}\right)$	$= \frac{1}{\sum\limits_{i=1}^{n} (\mathbf{X}_i)^2 - \left(\sum\limits_{i=1}^{n} \mathbf{X}_i\right)^2} \begin{bmatrix} n & -\sum\limits_{i=1}^{n} \mathbf{X}_i \\ -\sum\limits_{i=1}^{n} \mathbf{X}_i & \sum\limits_{i=1}^{n} (\mathbf{X}_i)^2 \end{bmatrix} \begin{pmatrix} \sum\limits_{i=1}^{n} \mathbf{X}_i \mathbf{Y}_i \\ \sum\limits_{i=1}^{n} \mathbf{Y}_i \end{pmatrix} .$	(6.13
Equation 6.1	3, may be written in a more conventional forms as	
	$a = \frac{n \sum_{i=n}^{n} (X_i Y_i) - \left(\sum_{i=1}^{n} X_i\right) \left(\sum_{i=1}^{n} Y_i\right)}{\frac{n}{2} (2 - 2) \left(\frac{n}{2} - 2\right)^2} ,$	(6.14

 \checkmark



$$b = \frac{\left(\sum_{i=1}^{n} \mathbf{Y}_{i}\right)\left(\sum_{i=1}^{n} \mathbf{X}_{i}^{2}\right) - \left(\sum_{i=1}^{n} \mathbf{X}_{i}\right)\left(\sum_{i=1}^{n} \mathbf{X}_{i}\mathbf{Y}_{i}\right)}{n\sum_{i=1}^{n} \mathbf{X}_{i}^{2} - \left(\sum_{i=1}^{n} \mathbf{X}_{i}\right)^{2}} \quad .$$
(6.15)

Equations 6.14 and 6.15 give values of coefficients a and b which best fit the given data set. Following the procedure used above, equations for constants corresponding to higher order curve fits can be determined.

6.3. General form of least squares

The method shown in Section 6.2 is specific to linear, or first order, least squares curve fits. It should be noted that any order polynomial can be fit to data, provided there are more data points than the order of the polynomial (i.e. four points are needed to perform a third order curve fit).

To fit n data points to a general m^{th} order polynomial given by

$$Y = A_1 X^m + A_2 X^{m-1} + \dots + A_m X + A_{m+1} \quad , \tag{6.16}$$

an expanded form of the matrix given by Eq. 6.11 is used. This matrix takes the general form of

$$\begin{bmatrix} \sum_{i=1}^{n} X_{i}^{2m} & \sum_{i=1}^{n} X_{i}^{2m-1} & \sum_{i=1}^{n} X_{i}^{2m-2} & \cdots & \sum_{i=1}^{n} X_{i}^{m} \\ \sum_{i=1}^{n} X_{i}^{2m-1} & \sum_{i=1}^{n} X_{i}^{2m-2} & \ddots & \ddots & \vdots \\ \sum_{i=1}^{n} X_{i}^{2m-2} & \ddots & \ddots & \sum_{i=1}^{n} X_{i}^{2} \\ \vdots & \ddots & \ddots & \sum_{i=1}^{n} X_{i}^{2} & \sum_{i=1}^{n} X_{i} \\ \sum_{i=1}^{n} X_{i}^{m} & \cdots & \sum_{i=1}^{n} X_{i}^{2} & \sum_{i=1}^{n} X_{i} & n \end{bmatrix} \begin{bmatrix} A_{1} \\ A_{2} \\ \vdots \\ A_{m} \\ A_{m+1} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} X_{i}^{m} Y_{i} \\ \vdots \\ \sum_{i=1}^{n} X_{i}^{m} Y_{i} \\ \vdots \\ \sum_{i=1}^{n} X_{i} Y_{i} \\ \sum_{i=1}^{n} Y_{i} \end{bmatrix} , \quad (6.17)$$

when considering the case of the polynomial given by Eq. 6.16.

To illustrate this general least squares approach, consider the 3rd order polynomial given by



 $Y = aX^3 + bX^2 + cX + d$

(6.18) n

To solve for the four coefficients a, b, c, and d, Eq. 6.17 is applied to Eq. 6.18 and results in

1		
	$\begin{bmatrix} \sum_{i=1}^{n} X_{i}^{6} & \sum_{i=1}^{n} X_{i}^{5} & \sum_{i=1}^{n} X_{i}^{4} & \sum_{i=1}^{n} X_{i}^{3} \\ \sum_{i=1}^{n} X_{i}^{5} & \sum_{i=1}^{n} X_{i}^{4} & \sum_{i=1}^{n} X_{i}^{3} & \sum_{i=1}^{n} X_{i}^{3} \\ \sum_{i=1}^{n} X_{i}^{4} & \sum_{i=1}^{n} X_{i}^{3} & \sum_{i=1}^{n} X_{i}^{2} & \sum_{i=1}^{n} X_{i} \\ \sum_{i=1}^{n} X_{i}^{3} & \sum_{i=1}^{n} X_{i}^{2} & \sum_{i=1}^{n} X_{i} & n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} X_{i}^{3} Y_{i} \\ \sum_{i=1}^{n} X_{i}^{2} Y_{i} \\ \sum_{i=1}^{n} X_{i}^{2} \sum_{i=1}^{n} X_{i} \\ \sum_{i=1}^{n} X_{i}^{2} & \sum_{i=1}^{n} X_{i} \\ \end{bmatrix} \cdot (6 $.19)

6.4. Example problems

6.4.1. Problem statement

A set of data, representing object's velocity measured as a function of time, is given in Table 6.5.

- 1) represent these graphically on a T-t (i.e., velocity versus time) diagram.
- determine mathematical formula, where V is velocity, t is time, and a and b are constants to be determined, for the data of Table 6.5, using the method of least squares. Show all work.
- 3) compute errors, E, by determining differences between measured and curve fitted points.
- 4) plot results of step 3 on an E-t diagram.
- 5) discuss your results.

Table 6.5. Data set representing object's velocity measured as a function of time.

t (sec)	V (ft/sec)					
0	1.45					
1	2.60					
2	3.60					
3	4.50					
4	4.95					
5	5.20					
6	5.10					
7	4.75					
8	4.10					





ring Departmer



and

$$a(36) + b(9) = (36.25)$$
 (6.22)

Solving Eqs. 6.21 and 6.22 for a and b results in values of a = 0.3458 and b = 2.6442. Using these values in Eq. 6.2 produces the equation for a linear least squares curve fit of the points in Table 6.5. This resultant equation is

$$Y = 0.3458 X + 2.6442$$
 (6.23)

Equation 6.23 can be used to compute the an fitted set of points corresponding to those listed in Table 6.5. These fitted points are given in Table 6.6.

Table 6.6. Data set representing object's velocity measured as a function of time shown with data from first order least squares curve fit.

t (sec)	V [data] (ft/sec)	V [fitted] (ft/sec)	Difference (ft/sec)	
0	1.45	2.64	-1.19	
1	2.60	2.99	-0.39	
2	3.60	3.34	0.26	
3	4.50	3.68	0.82	
4	4.95	4.03	0.92	
5	5.20	4.37	0.83	
6	5.10	4.72	0.38	
7	4.75	5.06	-0.31	
8	4.10	5.41	-1.31	

6.4.2. Problem statement

A set of data representing an object's velocity measured as a function of time, is given in Table 6.5.

- Determine mathematical formula V = V(t) = a t² + b t + c, where V is velocity, t is time, and a, b, and c are constants to be determined, for the data Table 6.1, using the method of least squares. Show all work.
- 2) Compute errors, E, by determining differences between measured and curve fitted points.
- 3) Plot results of step 3 on an E-t diagram.





The polynomial to which the data is to be fit is in the form of

$$Y = a X^2 + b X + c (6.25)$$

(6.27)

The three simultaneous equations to be used to fit the measured data to Eq. 6.25 are

$$a\sum_{i=1}^{9}X_{i}^{4} + b\sum_{i=1}^{9}X_{i}^{3} + c\sum_{i=1}^{9}X_{i}^{2} = \sum_{i=1}^{9}X_{i}^{2}Y$$
,

$$a \sum_{i=1}^{9} X_i^3 + b \sum_{i=1}^{9} X_i^2 + c \sum_{i=1}^{9} X_i = \sum_{i=1}^{9} X_i Y$$
, and (6.28)

$$a\sum_{i=1}^{9} X_{i}^{2} + b\sum_{i=1}^{9} X_{i} + c\sum_{i=1}^{9} 1 = \sum_{i=1}^{9} .$$
 (6.29)

By putting the values from Eqs. 6.20 and 6.24 into the Eqs. 6.27 to 6.29, the relations

$$a(8772) + b(1296) + c(204) = 945.45$$
, (6.30)

$$a(1296) + b(204) + c(36) = 165.75$$
, and (6.31)

$$a(204) + b(36) + c(9) = 36.25$$
 (6.32)

Solving Eqs. 6.30 to 6.32 for *a*, *b*, and *c*, the values a = -0.13706, b = 1.442370, and c = 1.365151 are obtained. Substituting these values into Eq. 6.25, the equation representing the second order curve fit of the points given in Table 6.5 is obtained. This equation is

$$Y = -0.13706 X^2 + 1.44237 X + 1.36515 .$$
(6.33)

Equation 6.33 can be used to compute the an fitted set of points corresponding to those listed in Table 6.5. These fitted points are given in Table 6.7.

Table 6.7. Data set representing object's velocity measured as a function of time shown with data from second order least squares curve fit.

t	V [data]	V [fitted]	Difference
(sec)	(ft/sec)	(ft/sec)	(ft/sec)
0	1.45	1.37	0.08
1	2.60	2.67	-0.07
2	3.60	3.70	-0.10
3	4.50	4.46	0.04
4	4.95	4.94	0.01
5	5.20	5.15	0.05
6	5.10	5.09	0.01
7	4.75	4.75	0.00
8	4.10	4.13	-0.03



Note that the differences shown in Table 6.7 and substantially smaller than those presented in Table 6.6. Also notice that the data has a somewhat random appearance in Fig. 6.8 while in Fig. 6.7 the data follows a distinct pattern. All this evidence points to the conclusion that a second order curve fit of the data is far better than a first order curve fit. 1.5 1 VELOCITY, ft/sec 0.5 0 -0.5 -1 -1.5 2 0 4 6 8 TIME, sec Fig. 6.8. Graphical representation of the data in Table 6.7 showing the difference between the measured points and the curve fitted points.

References

Anscombe, F. J., "Graphs in statistical analysis," American Statistician, 27:17-21, 1973. Pryputniewicz, R. J., "Engineering Experimentation," Worcester Polytechnic Institute, 1993.





Curve fitting: matrix approach Discussed in class: estimation of \pm ranges for fitting coefficients

Curve fitting can be expressed as:

 $[X'X]\beta = (X'Y) \qquad ([X'] \text{ is transpose of } [X])$

Obtained from model: $[X]\beta = (Y)$ (Over determined system of equations)

Vector of unknowns (coefficients of model used, e.g., slope and dc values in a linear fit)

Solution is: $\beta = [X'X]^{-1}(X'Y)$

Components of inverse matrix are expressed as:

$$[X'X]^{-1} = [c_{ij}]$$





Curve fitting: matrix approach Discussed in class: estimation of \pm ranges for fitting coefficients

Recall that sum of the square of the errors (SCE) is:





Curve fitting: matrix approach Discussed in class: estimation of \pm ranges for fitting coefficients

Confidence interval $(1 - \alpha)$ 100% for fitting coefficient β_i :



Interval for β_i fitting coefficients where their values are $(1 - \alpha)100\%$ reliable





Curve fitting: matrix approach

Discussed in class: estimation of \pm ranges for fitting coefficients

TABLE of CRITICAL VALUES for STUDENT'S t DISTRIBUTIONS

degrees of freedom

U				Column	headings	denote p	robabilitie	es (α) abo	ove tabula	ated value	es.			
Ĩ	df	0.40	0.25	0.10	0.05	0.04	0.025	0.02	0.01	0.005	0.0025	0.001	0 0005	
-	1	0.325	1.000	3.078	6.314	7.916	12,706	15.894	31.821	63.656	127.321	318,289	636.578	
	2	0.289	0.816	1.886	2.920	3.320	4,303	4.849	6.965	9.925	14.089	22.328	31,600	
	3	0.277	0.765	1.638	2.353	2.605	3.182	3.482	4.541	5.841	7.453	10.214	12.924	
	4	0.271	0.741	1.533	2.132	2.333	2.776	2.999	3.747	4.604	5.598	7.173	8.610	1 Student
	5	0.267	0.727	1.476	2.015	2.191	2.571	2.757	3.365	4.032	4.773	5.894	6.869	
	6	0.265	0.718	1.440	1.943	2.104	2.447	2.612	3.143	3.707	4.317	5.208	5.959	
	7	0.263	0.711	1.415	1.895	2.046	2.365	2.517	2.998	3.499	4.029	4.785	5.408) baae 114 o
	8	0.262	0.706	1.397	1.860	2.004	2.306	2.449	2.896	3.355	3.833	4.501	5.041	
	9	0.261	0.703	1.383	1.833	1.973	2.262	2.398	2.821	3.250	3.690	4.297	4.781	
	10	0.260	0.700	1.372	1.812	1.948	2.228	2.359	2.764	3.169	3.581	4.144	4.587	
	11	0.260	0.697	1.363	1.796	1.928	2.201	2.328	2.718	3.106	3.497	4.025	4.437	
	12	0.259	0.695	1.356	1.782	1.912	2.179	2.303	2.681	3.055	3.428	3.930	4.318	
	13	0.259	0.694	1.350	1.771	1.899	2.160	2.282	2.650	3.012	3.372	3.852	4.221	
	14	0.258	0.692	1.345	1.761	1.887	2.145	2.264	2.624	2.977	3.326	3.787	4.140	
	15	0.258	0.691	1.341	1.753	1.878	2.131	2.249	2.602	2.947	3.286	3.733	4.073	Vou can a
	16	0.258	0.690	1.337	1.746	1.869	2.120	2.235	2.583	2.921	3.252	3.686	4.015	, <i>, , , , , , , , , , , , , , , , , , </i>
	1/	0.257	0.689	1.333	1.740	1.862	2.110	2.224	2.507	2.898	3.ZZZ	3.640	3.905	
	10	0.257	0.000	1.000	1.734	1.000	2.101	2.214	2.002	2.0/0	3.197	3.010	3.922	TO ANAIVSE
	19	0.257	0.000	1.320	1.729	1.000	2.095	2.205	2.009	2.00	3.174	3.579	3.003	
	20	0.257	0.007	1.320	1.720	1.044	2.000	2.197	2.020	2.040	3.105	3.552	3,000	I lab
	22	0.256	0.000	1.323	1.721	1.835	2.000	2.105	2.510	2.03	3 110	3.505	3 702	
	23	0.256	0.685	1.319	1 714	1.832	2.069	2.103	2.500	2.013	3 104	3 485	3 768	
	24	0.256	0.685	1.318	1711	1.828	2.064	2 172	2 492	2 797	3.091	3 467	3 745	
	25	0.256	0.684	1.316	1.708	1.825	2.060	2.167	2.485	2.787	3.078	3,450	3,725	
	26	0.256	0.684	1.315	1.706	1.822	2.056	2,162	2.479	2,779	3.067	3.435	3,707	
	27	0.256	0.684	1.314	1.703	1.819	2.052	2.158	2.473	2.771	3.057	3.421	3.689	
	28	0.256	0.683	1.313	1.701	1.817	2.048	2.154	2.467	2.763	3.047	3.408	3.674	
	29	0.256	0.683	1.311	1.699	1.814	2.045	2.150	2.462	2.756	3.038	3.396	3.660	1
	30	0.256	0.683	1.310	1.697	1.812	2.042	2.147	2.457	2.750	3.030	3.385	3.646	
	31	0.256	0.682	1.309	1.696	1.810	2.040	2.144	2.453	2.744	3.022	3.375	3.633	
	32	0.255	0.682	1.309	1.694	1.808	2.037	2.141	2.449	2.738	3.015	3.365	3.622	
	33	0.255	0.682	1.308	1.692	1.806	2.035	2.138	2.445	2.733	3.008	3.356	3.611	
	34	0.255	0.682	1.307	1.691	1.805	2.032	2.136	2.441	2.728	3.002	3.348	3.601	
	35	0.255	0.682	1.306	1.690	1.803	2.030	2.133	2.438	2.724	2.996	3.340	3.591	
	36	0.255	0.681	1.306	1.688	1.802	2.028	2.131	2.434	2.719	2.990	3.333	3.582	
	37	0.255	0.681	1.305	1.687	1.800	2.026	2.129	2.431	2.715	2.985	3.326	3.574	
	38	0.255	0.681	1.304	1.686	1.799	2.024	2.127	2.429	2.712	2.980	3.319	3.566	
	39	0.255	0.681	1.304	1.685	1.798	2.023	2.125	2.426	2.708	2.976	3.313	3.558	
	40	0.255	0.681	1.303	1.684	1.796	2.021	2.123	2.423	2.704	2.9/1	3.307	3.551	
	60	0.254	0.679	1.296	1.6/1	1.781	2.000	2.099	2.390	2.000	2.915	3.232	3.460	
	100	0.254	0.677	1.292	1.660	1.770	1.990	2.000	2.3/4	2.039	2.007	2 174	3.410	
	120	0.254	0.677	1.290	1,659	1.766	1.904	2.00	2.304	2.020	2.07	3.160	3 373	
	140	0.254	0.676	1.205	1.656	1 763	1.000	2.073	2.000	2.017	2.000	3 149	3 361	
	160	0.254	0.676	1.200	1.654	1 762	1.975	2.071	2.350	2.607	2.847	3 142	3 352	
	180	0.254	0.676	1.286	1.653	1.761	1.973	2.069	2.347	2.603	2.842	3.136	3.345	
	200	0.254	0.676	1,286	1.653	1,760	1.972	2.067	2 345	2.601	2 838	3 131	3 340	1
	250	0.254	0.675	1.285	1.651	1,758	1.969	2.065	2.341	2,596	2.832	3,123	3,330	
	inf	0.253	0.674	1.282	1.645	1.751	1.960	2.054	2.326	2.576	2.807	3.090	3.290	cai Engineering Departmer
		0.200	0.014	1.202	1.040	1.791	1.000	2.004	2.020	2.010	2.007	0.000	0.200	

Student's t - see page 114 of Holman

You can apply this to analyses done in Lab #2

