Appendix A

Guidelines for Data Approximation

In this context, data approximation refers to the determination of a continuous function which represents a set of discrete data in a satisfactory manner. The choice of a function and the judgment as to what is satisfactory will depend upon the application, and consequently FC offers a repertoire of methods suitable for a broad range of problems.

The principal part of this appendix is a selection table which correlates the available methods with the parameters which characterize the application. Following this table there is a set of notes which explain the table entries and some examples of its use. It should be possible to identify one or more methods which are compatible with almost any application. However, it is to be expected that some problems will not lend themselves to the specific characterization of the table. In such cases, some experimentation may be required. A further alternative is to modify the problem in order to mold it to the available tools, by gathering more (or better) data and/or by preprocessing the data through filtering or smoothing.

It should be noted that the table entries reflect typical conditions rather than absolute constraints. It is likely that some problems can be solved by apparently disqualified methods, but it is usually preferable to try the recommended methods first. A final preliminary caution concerns what to do if there appears to be insufficient information about the problem to use the table constructively. One approach, which is not recommended, is simply to assume the worst and choose a technique which handles the broadest scope of problems. Instead, and in fact in all instances where it is practical, the data should be analysed before attempting an approximation. For one-dimensional problems, an x-y plot is useful, and even for surface fits, contour plots can serve a similar purpose.

Method Selection Table

Each row of the selection table prescribes a characteristic parameter of the approximation problem. Not all will apply in every case, since the list is fairly comprehensive. The significance of each parameter is explained in the notes following the table.

The columns refer to the approximation methods presently available according to the following list:

Column	Name	Class	Section
1	POLYREG	UTILITY	3.3.1
2	CHEBFIT	UTILITY	3.3.5
3	POLYDEG	UTILITY	3.3.3
4	POLYEPS	UTILITY	3.3.3
5	FITxxx	UTILITY	3.3.3
6	SFIT	UTILITY	3.3.4
7	AJAX	EQUATIONS SOLVER	4.2

CHARACTERISTIC	OPTIONS	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Approximation Norm	L1							
	L2	X		X	X	X	X	X
	Minimax		X					
Approximation Rule	Minimum Norm	X	X	X		X	X	X
	Best Minimum Norm	X	(1)	(1)				
	Bounded Norm	(1)	(1)		X	X	X	
Function Form	Polynomial	X	X	X	X	X	X	X
	General Linear					X	X	X
	General Nonlinear							X
	Phenomological							X
Function Parameters	Medium (8 to 20)	(2)	(2)	(2)	X	X	X	
	Many (more than 20)							X
Parameter Estimates	Available	X	X	X	X	X	X	X
	Unavailable/Poor	X	X	X	X	X	X	
Function Constraints	None	X	X	X	X	X	X	X
	Fixed limits							X
	Linear							(3)
	Nonlinear							(3)
Data Quantity	Few (to 100)	X	X	X	X	X	X	X
	Medium (to 1000)	X	X	X	X	X	X	
	Many	X	X					
Data Distribution	Regular	X	X	X	X	X	X	X
	Irregular	(4)	X	(4)	(4)	(4)	(4)	(4)
Data Precision	High	X	X	X	X	X	X	X
	Low	X		X	X	X	X	X
	Variable					X	X	
Data Dimensionality	N > 1						X	X
Efficiency Needed	High	X	X					
	Medium	X	X	X	X	X	X	

Column	Name	Class	Section
8	MARS	EQUATIONS SOLVER	4.3
9	HERA	OPTIMIZATION SOLVER	6.1
10	HERCULES	OPTIMIZATION SOLVER	7.5
11	THOR	OPTIMIZATION SOLVER	7.3
12	JUPITER	OPTIMIZATION SOLVER	7.4
13	ZEUS	OPTIMIZATION SOLVER	7.2
14	JOVE	OPTIMIZATION SOLVER	7.1

CHARACTERISTIC	OPTIONS	[8]	[9]	[10]	[11]	[12]	[13]	[14]
Approximation Norm	L1		X	X	X	X	X	X
	L2	X	X		X	X	X	X
	Minimax							
Approximation Rule	Minimum Norm	X	X	X	X	X	X	X
	Best Minimum Norm							
	Bounded Norm							
Function Form	Polynomial	X	X	X	X	X	X	X
	General Linear	X	X	X	X	X	X	X
	General Nonlinear	X	X		X	X	X	X
	Phenomenological	X	X	X	X	X	X	X
Function Parameters	Medium (8 to 20)	X	X	X	X	X	X	X
	Many (more than 20)	X		X	X	X	X	
Parameter Estimates	Available	X	X	X	X	X	X	X
	Unavailable/Poor			X				
Function Constraints	None	X	X			X	X	X
	Fixed limits	X		X	X	X	X	X
	Linear	(3)		X	X	X	X	X
	Nonlinear	(3)			X	X	X	X
Data Quantity	Few (to 100)	X	X	X	X	X	X	X
	Medium (to 1000)		X	X	X	X	X	X
	Many		X	X	X	X	X	X
Data Distribution	Regular	X	X	X	X	X	X	X
	Irregular	(4)		(4)				
Data Precision	High	X	X	X	X	X	X	X
	Low	X	X	X	X	X	X	X
	Variable		X	X	X	X	X	X
Data Dimensionality	N>1	X	X		X	X	X	X
Efficiency Needed	High							
	Medium							

Entries in the table are coded in the following manner:

Blank : Method unqualified or not recommended

X : Method qualified

(number) : Method conditionally qualified (refer to number note)

Note that there are additional factors which discriminate among the methods, involving special capabilities, convenience features, report generation, and usage limitations. Consult the referenced sections for further details.

Table Notes

(1) The rule is not automated for this method but it is easily accomplished by a simple loop and test logic in the calling program. For example, the use a bounded norm with CHEBFIT:

```
      DO 10 NDEG=LITTLE,LARGE
        @CHEBFIT(NPTS,X,Y,NDEG,ERR,A,B,0)
        IF(ERR.LT.EPSILON) GOTO 20
10    CONTINUE
20    CONTINUE
```

(2) These methods are doubtful in the upper range of values

(3) Inequality constraints require the use of slack variables

(4) Only use low order polynomials (< 4) and mildly nonlinear functions.

Selection Table Terminology

Approximation Norm - This the measure used to evaluate the quality of the approximation. Defining the deviation to be the absolute difference between the function and the data, the three most common norms are:

L1 - The sum of the deviations

L2 - The sum of squares of the deviations

Minimax - The maximum deviation

The method will attempt to minimize the norm. For the L2 norm, this is the familiar case of least-squares approximation. Each norm has a domain of preferred applications, but unless guided otherwise, the L2 norm is generally preferable. This is because, if the data errors are uniform and normally distributed with zero mean, the maximum liklihood estimator is given by a least-squares approximation. An L1 norm is useful when the data precision is variable or when wild points are suspected. The minimax norm is advantageous when the data have low errors but are irregularly distributed over the approximation interval.

Approximation Rule - This is the criterion for performing the approximation. The basic rule is simply to minimize the chosen approximation norm by an appropriate selection of the function parameters. Some methods allow the function to have a variable number of undetermined parameters (see function form and parameters, below). In this case, the best minimum norm can be calculated by trying different numbers of parameters within the allowable range. The technique usually starts with the simplest function and adds elements one by one while comparing the results. All combinations may not be tried if the method can determine an effective enhancement sequence. Alternatively, a maximum permissible error can be specified, in which case the trials continue until the norm satisfies the prescribed bound. The bounded norm rule has two major advantages. First, the quality of fit is controlled, rather than being dependent upon a clever or fortunate choice of a good functional form. Secondly, most methods will choose the simplest function which satisfies the bound. This can help minimize numerical problems and computational complexity in subsequent use of the function.

Function Form and Parameters - The approximation function must be sufficiently continuous, which means that there can be no discontinuities at the discrete values of the independent variable(s) for which the data are evaluated. It must also contain one or more undetermined parameters which will be calculated by the method to satisfy the approximation rule. The function must be continuous in these parameters, at least over their allowable range. The classes of functions may be categorized as follows. (The examples assume one independent variable, X, and the parameters, A, B, C, ...)

(1) General functions, nonlinear in the parameters

A*EXP(B*X)   SIN(A*X)*(B+C*X)   X**A

(2) General functions, linear in the parameters

A*EXP(X)+B*EXP(X)   A*X+B*X**3

(3) Special case of (2), a polynomial

A+B*X+C*X**2

(4) Phenomenological, where the functions are general nonlinear ones (including differential equations) derived from the physical laws govering the phenomena being measured. This class of functions has the advantage that the functions will faithfully predict discontinuities and asymptotic behavior in specific phenomena regimes, even when the data does not necessarily suggest this behavior on a general basis. (Applications 8-1 and 8-3 are good examples of phenomenological fitting.)

Some methods allow the function to be expressed as a series containing a variable number of terms, each with one parameter. The method will then determine both the parameter values and the appropriate number of terms to be used. The simplest case is a power series in X, i.e. a general polynomial. An example is POLYEPS (FIT) which can compute a bounded norm for a general polynomial function. Where the method allows more complex functions, the terms of the series can be inhomogeneous, i.e. the terms may have no common functional character. (An example of a homogeneous general series is a Fourier series.) Finally, some methods require initial estimates for the parameters, usually good ones if success is to be assured.

Function Constraints - It is sometimes necessary, on mathematical or physical grounds, to limit the allowable range of parameter values. The simplest constraints are fixed limits, i.e. constant inequality constraints such as

$$A \geq 0 \quad 5 \leq B \leq 10$$

Another class consists of linear equality or inequality constraints such as

$$A = B + C \quad A \geq (B+10)\,.$$

And a final class consists of nonlinear equality or inequality constraints such as

$$A^*(1+A) = EXP(B) \quad A \geq SQRT(B)\,.$$

Data characteristics - The data distribution is regular when the measurements are spread fairly uniformly over the approximation interval. It is irregular when there are conspicuously sparse or dense regions. Weighted norms are usually necessary for irregular data.

Data precision is measured by the expected error, e.g. by the standard deviation if the errors are uniformly distributed. The judgment as to what constitutes high precision is a little subjective. Variable precision data, i.e. those subject to systematic errors or containing wild points, always require weighting and are best handled by suitable preprocessing.

Dimensionality refers to the number of independent variables. One-dimensional problems have data expressed as a bilinear table and the approximation is a curve fit. For two or more dimensions, a surface fit is involved.

Efficiency - This factor is very crude. Efficiency becomes a significant factor in an application when some combination of the following conditions occur:

(1) There are many data

(2) The form and conditions of the approximation make it difficult.

(3) The approximation is an integral part of a larger problem and/or it must be repeatedly performed.

Examples of Method Selection

(1) Data characteristics:

Number	500
Distribution	Regular
Precision	Low
Dimensionality	1

Approximation requirements

Norm	Least-squares
Rule	Accuracy of 0.0001 (bounded norm)
Function	Polynomial
Constraints	None
Parameters	Unknown

The qualified methods are POLYEPS, FIT, SFIT, and POLYREG. If efficiency is crucial, then POLYREG is best.

(2) Data characteristics:

Number	50
Distribution	Regular
Precision	High
Dimensionality	2

Approximation requirements

Norm	L1
Rule	Minimum norm
Function	General linear
Constraints	Fixed limits
Parameters	3 with good estimates

The qualified methods are THOR, JUPITER, ZEUS, and JOVE. If efficiency is crucial, then a satisfactory result might be obtained using a least squares without constraints by means of SFIT. (Often fixed limits play no role in the approximation.)