Modeling of Data

- *NRiC* Chapter 15.
- Model depends on adjustable parameters.
- Can be used for "constrained interpolation".
- Basic approach:

1. Choose *figure-of-merit* function (e.g. χ^2).
2. Adjust *best-fit parameters*: minimize merit function.
3. Compute *error estimates* for parameters.
4. Compute *goodness-of-fit*.

Least Squares Fitting

- Suppose we want to fit N data points (x_i,y_i) with a function that depends on M parameters a_j and that each data point has a standard deviation σ_i. The *maximum likelihood estimate* of the model parameters is obtained by minimizing:

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - y(x_i; a_1 \quad a_M)}{\sigma_i} \right)^2$$

- Assuming the errors are normally distributed, a "good fit" has $\chi^2 \sim \nu$, where $\nu = N - M$.

Fitting Data to a Straight Line

- For this case the model is simply:

  $y(x) = y(x; a, b) = a + bx$

- Derive formula for best-fit parameters by setting $\partial \chi^2/\partial a = 0 = \partial \chi^2/\partial b$.

- Derive uncertainties in a and b using:

$$\sigma_f^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial f}{\partial y_i} \right)^2$$

- Want Q = `gammq`$((N - 2)/2, \chi^2/2) > 0.001$.

General Linear Least Squares

- Can generalize to any linear combination:

$$y(x) = \sum_{j=1}^{M} a_j X_j(x)$$

e.g. $y(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_M x^{M-1}$.

- Define $N \times M$ *design matrix* $A_{ij} = X_j(x_i)/\sigma_i$.
- Also define vector **b** of length N: $b_i = y_i/\sigma_i$ and vector **a** of length M: $a_i = a_1,\ldots,a_M$.
- Then we wish to find **a** that minimizes:

$\chi^2 = |A\mathbf{a} - \mathbf{b}|^2 \quad \leftarrow \quad$ This is what SVD solves!

General Linear Least Squares, Cont'd

- Recall for SVD we had $A = UWV^T$.
- Rewriting the SVD solution we get:

$$\mathbf{a} = \sum_{j=1}^{M} \left(\frac{U_{(j)}\mathbf{b}}{w_j} \right) V_{(j)}$$

where $U_{(j)}$ and $V_{(j)}$ denote columns of U and V.

- As before, if w_j is small (or zero), can omit.
  - Useful because least-squares problems are *both* overdetermined ($N > M$) *and* underdetermined (ambiguous combinations of parameters exist)!

Nonlinear Models

- Suppose model depends *nonlinearly* on the a_j's...

  e.g. $y(x) = a_1 \exp(-a_2 x^2)$.

- Still define χ^2, but must proceed iteratively:
  - Use $\mathbf{a}_{next} = \mathbf{a}_{cur} - \lambda \nabla \chi^2(\mathbf{a}_{\mathbf{cur}})$ far from minimum (steepest descent), where λ is a constant.
  - Use $\mathbf{a}_{\mathbf{next}} = \mathbf{a}_{\mathbf{cur}} - D^{-1}[\nabla \chi^2(\mathbf{a}_{\mathbf{cur}})]$ close to minimum, where D is the *Hessian* matrix.
  - The *Levenberg-Marquardt method* adjusts λ to smooth the transition between these two regimes.

Levenberg-Marquardt Method

- *NRiC* provides two routines, `mrqmin()` and `mrqcof()`, that implement the L-M method.
- The user must provide a function that computes $y(x_i)$ as well as all the partial derivatives $\partial y/\partial a_j$ evaluated at x_i.
- The routine `mrqmin()` is called iteratively until a successful step (i.e. one in which λ gets smaller) changes χ^2 by less than a fractional amount, like 0.001 (no point in doing better).