Project 1: One Compartment Modeling

- The most commonly employed approach to the pharmacokinetics characterization of a drug
- Represent the body as a system of compartments
- Compartments usually have no physiological or anatomic reality
- Assume that the rate of transfer between compartments and the rate of drug elimination from compartments follows first order or linear kinetics
- One compartment = the body = a single, kinetically homogeneous unit
- Assume that the rate of change of drug concentration in plasma reflects quantitatively the change in drug concentrations through the body
- The elimination of most drugs in humans and animals following therapeutic or nontoxic dose can be characterized as an apparent first-order process (i.e., the rate of elimination of drug from the body at any time is proportional to the amount of drug in the body at that time.)

1. Intravenous Injection (drug concentration in the plasma)

- Following rapid intravenous injection of a drug that distributes in the body according to a one-compartment model and is eliminated by apparent first-order kinetics.
- The rate of loss is given by $dx/dt = -kx$

where x = the amount of drug in the body, k=the apparent first-order elimination rate constant for the drug, (-) = the negative sign indicates that drug is being lost from the body.

- By the Laplace transform

$$s\overline{X} - X_0 = -k\overline{X}$$

$$\overline{X} = X_0/(s+k)$$

$$x = x_0 e^{-kt}$$

- Taking the natural logarithm of both sides gives

$$\ln x = \ln x_0 - kt$$

Since $2.303 \log a = \ln a$

$$\log x = \log x_0 - \frac{kt}{2.303}$$

If the volume of the compartment is known

$$X = VC \text{ or } C = X / V$$

where C=drug concentration

$$\log C = \log C_0 - \frac{kt}{2.303}$$

C_0 = the drug concentration in plasma immediately after injection.

2. Simulation

- You know all the details of a given model or system
- Does not require data
- Allows specific testing of a known (or unknown) system
- Example: Monoexponential

Let's say we performed a drug study and determined that the time course of a given drug following a bolus injection can be described by the following simple exponential equation.

$$y = 75e^{-kt}$$

Let's also say that we know for healthy subjects, k is equal to 0.05.

(Experiment 1)

Plot y(t) for t is from 0 to 120 minutes with k=0.05 at every 0.5min. Label x and y-axis.

Now suppose we would like to know how the time course for this drug would change if the rate of disappearance were to change ±2-fold. The change in k might reflect changes due to renal dysfunction or changes in the degradation of the drug in vivo. We would use simulation to calculate the new time course without having to do a new experiment.

(Experiment 2)

Plot t vs. y(t) for k=0.05, 0.025, and 0.10. Use different line formats.

3. Parameter Identification

i. Collect data
ii. Test a hypothesized model
iii. Don't know details of the model or system

Let's take our example from above. We took our drug and performed experiments in which we injected some known amount into an experimental subject and measured the concentration over time. Our data might look like this…

How do we determine the value of k? We can use the method of least squares to find k from the data we have collected

4. Least Squares Method

Let's use the linear regression as an example. We want to draw a straight-line which best "fits" or represents some data we have collected. What determines the best line through the data? The best line is that which comes as close as to the data as possible, while having the data randomly distributed about the time.

The least squares method works by minimizing the distance between our measured data and that predicted by our equation. The distance between the measured data and the predicted value is known as the residual and is defined.

$$\text{Residual} = Y_i - \hat{Y}_i$$

The residuals are squared and totaled to calculate the residual sum-of-squares (SSQ)

$$SSQ = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Least squares work by basically testing out various combinations of parameters to determine which provides the lowest SSQ. The SSQ is known as a parameter of "goodness-of-fit" and can be used in various statistical analyses.

The method of least squares can be applied to both linear and non-linear problems. Linear least squares problems can be solved algebraically.

4-1. Linear Least Squares

1. By hands

(Experiment 3)

Find k from the linearized data computed in Experiment 1. Discuss the results. Any errors in k?

2. Fitting a line

Given $y = a + bx$

Fit a line through given points (x_1,y_1), ..., (x_n, y_n) so that the sum of the squares of the differences of those points from the straight line is minimum

Minimize q (or SSQ) depends on a and b, thus a necessary condition for q to be minimum is

$$\frac{dq}{da} = -2\sum(y_i - a - bx_i) = 0$$
$$\frac{dq}{db} = -2\sum x_i(y_i - a - bx_i) = 0$$

Rewrite

$$an + b\sum x_i = \sum y_i$$
$$a\sum x_i + b\sum {x_i}^2 = \sum x_i y_i$$

In matrix form

$$\begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum {x_i}^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

3. Fitting a function (generalized form)

Let the data points be x_i, y_i where i=1,2,...,m. The function to be fitted to the data has the form

$$y = f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \ldots + a_n\phi_n(x)$$

where $\phi_i(x)$ are chosen functions and a_i are unknown coefficients. Let the error between this function and the data point (x_k, y_k) be q_k. Thus

$$q_k = y_k - \{a_1\phi_1(x_k) + a_2\phi_2(x_k) + \ldots + a_n\phi_n(x_k)\}$$

Denoting the sum of the squares of these errors by S we have

$$S = \sum_{i=1}^{m}[y_i - \{a_1\phi_1(x_i) + a_2\phi_2(x_i) + \ldots + a_n\phi_n(x_i)\}]^2$$

We wish to minimize S and to do this we make

$$\frac{dS}{da_k} = 0, k = 1,2,\ldots,n$$

Now

$$\frac{dS}{da_k} = \sum_{i=1}^{m} 2[y_i - \{a_1\phi_1(x_i) + a_2\phi_2(x_i) + \ldots + a_n\phi_n(x_i)\}]\phi_k(x_i) = 0$$

Hence

$$\sum_{i=1}^{m} y_i\phi_k(x_i) - a_1\sum_{i=1}^{m}\phi_1(x_i)\phi_k(x_i) + a_2\sum_{i=1}^{m}\phi_2(x_i)\phi_k(x_i) + \ldots + a_n\sum_{i=1}^{m}\phi_n(x_i)\phi_k(x_i) = 0,$$
$$k = 1,2,\ldots,n$$

Rearranging these equations into matrix notation we have

$$\begin{bmatrix} \sum_{i=1}^{m}\{\phi_1(x_i)\}^2 & \sum_{i=1}^{m}\phi_2(x_i)\phi_1(x_i) & \ldots & \sum_{i=1}^{m}\phi_n(x_i)\phi_1(x_i) \\ \sum_{i=1}^{m}\phi_1(x_i)\phi_2(x_i) & \sum_{i=1}^{m}\{\phi_2(x_i)\}^2 & \ldots & \sum_{i=1}^{m}\phi_n(x_i)\phi_2(x_i) \\ \ldots & \ldots & \ldots & \ldots \\ \sum_{i=1}^{m}\phi_1(x_i)\phi_n(x_i) & \sum_{i=1}^{m}\phi_2(x_i)\phi_n(x_i) & \ldots & \sum_{i=1}^{m}\{\phi_n(x_i)\}^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} y_i\phi_1(x_i) \\ \sum_{i=1}^{m} y_i\phi_2(x_i) \\ \ldots \\ \sum_{i=1}^{m} y_i\phi_n(x_i) \end{bmatrix}$$

Or in matrix form

$$Pa = b \quad (*)$$

where

$$p_{ki} = \sum_{i=1}^{m}\phi_k(x_i)\phi_i(x_i) \text{ and } b_k = \sum_{i=1}^{m} y_i\phi_k(x_i)$$

(Experiment 4)

Derive $Pa = b$ by hands, a and b for $y(t) = b_0 + b_1x + b_2x^2$ according to the procedures in 2 and 3.

(Experiment 5)

Find k from the data in Experiment 1, by writing a Matlab program for linear fitting (i.e., use Eq. (*) and inv function of Matlab) Any errors of k?

4-2. Non-linear Least Squares

In non-linear least squares problems, the complexity of the system requires some algorithms to search for the solution to the problem. With non-linear least squares, an iterative process is used to estimate the parameters. There are many different non-linear least squares algorithms, but the most common is the Marquardt-Levenburg. The various algorithms differ only in the way they search for the best parameters.

(Experiment 6)

This time, generate noisy y(t) by adding some noise. Generate random noise using rand or randi. Noise should be between -0.5~0.5 or -1~1. Add noise to y(t). The amount of noise can be controlled by multiplying a small number (i.e., 0.1 or 5). After adding noise, you should apply abs to make positive concentration of drug.

Plot t vs. noise-free and t vs. noisy data for three different amounts of noise. Use k=0.05.

(Experiment 7)

Find k from the noise-free data from Experiment 1 and noisy data from Experiment 7 by writing a Matlab program for non-linear least squares fitting. (1) Plot SSQ vs. iteration, (2) Plot k_estimated vs. iteration, and (3) Plot fitting curve for every iteration. (Hint: use fminsearch.m and design a least square cost function.). Examine the estimated k value with respect to the noise level. Compute errors of k. Discuss the results.

5. Goodness-of-Fit

Goodness-of-fit assessment is not restricted to how well a given set of parameters fit the data. Goodness-of-fit also used to compare the ability of different models to account for the observed data. For example, in our drug study would a two-exponential function fit the data better? How do we show that the fit is improved with a different set of parameters or a different model? There are several parameters and statistics one can look at

Sums-of-Squares (SSQ)

The residual (or weighted) SSQ is a very good general indicator of goodness of fit. Recall that the least squares method minimizes SSQ to determine the best fit. One can compare SSQ (preferably weighted) to check for improvements in fits. CAUTION: The SSQ can be used to compare different models only if those models have the same number of parameters.

R-Square

R^2 indicates what proportion of the y-variable is accounted for by the x-variable. The higher the R^2, the better the model fit. The weakness of the R^2 as an indicator of goodness-of-fit is its relatively low sensitivity. R^2 can be high even if the fit is not good by alternative criteria. CAUSION: This statistic should not be used to compare different models.

Residual Plots

The SSQ and R^2 provide information on the overall fit of the data, but do not provide any goodness-of-fit information over the course of the data collected. For our drug study example, we can plot the residuals over time to look for systematic deviations from the measured data.

Ideally, the residuals should be randomly distributed about the zero line. Patterns in the residual plot typically indicate an inadequacy in the model structure, i.e., the model is not accounting for a process which contributes significantly to the observed dynamics. One can test for patterns in the residuals using a runs test, autocorrelation, or some similar methods. Residual testing is a powerful method for assessing differences in model fits. In our example, we have plotted the pure residuals ($Y - \hat{Y}$), but it is also common to plot residuals as a fraction (percent of the observed value), standardized to the mean data, or standardized to the known variance.

6. Other Terms

Error

This is the standard deviation (SD) in the parameter estimate. i.e., an indicator of the precision with which the parameter was estimated. In the literature, this number is many times reported as a percent or fractional SD (FSD).

Dependency

Dependency means that a change in any one parameter of the group can be (nearly) compensated for by changes in the other parameters of the group. If these values are all zero, none of the parameters in the model are dependent upon each other. The magnitude in this number indicates the level of dependency between parameters

Iterations

Indicates how many iterations were performed before the program quit. If convergence is

achieved, the program so notes. If convergence is not achieved, the program will also indicate as such.

Best Weighted Sum of Squares

This is the final estimate of the SSQ

Weighted Root Mean Square Error

This is the average standard deviation about zero calculated as follows:

$$RMS = \sqrt{\frac{SSQ}{n-p}}$$

where n is the number of data points, and p is the number of parameters estimated

Initial Parameter Estimates

Provided to the program as some guess or initial estimates about the parameter values. For each problem, there is a so-called parameter search space. Depending upon what the search space looks like, one runs the risk of hitting a local minimum. Therefore, it is important to choose initial parameter estimates which are physiologically plausible. For instance, one would not choose a plasma volume of 500 liters as an initial estimate.

Convergence Criteria

How do you decide when you have converged? After some time, the change in parameter estimates will have little to no effect on the fits to data. i.e., the fists and SSQ therefore do not change. Most routines will allow you to set a convergence criterion as a lower bound for a change in the SSQ. For example, one can state a priori that if the SSQ does not change by 10%, then convergence is achieved. Remember that each problem is unique and the search space could contain several local minima. Therefore, it is wise to choose a convergence criterion, e.g., 0.1% change in SSQ

Constraints

Suppose you know a priori that a certain parameter must be within a certain range of values. Let's take our drug study. The parameter k represents the rate of disappearance of the drug from the system, hence the negative sign in the exponent. If k were estimated to be a negative number, this would make the exponent positive, and therefore increase the drug over time. This would clearly be incorrect. To prevent such unreasonable parameter values, one can introduce constraints which indicate the range of values permissible for the

identification.

Weighting

All biological data are imprecise. There are errors due to data collection and assay which result in noise in the data. The more imprecise the original measurement, the more noise in your data, and the less precise your parameter estimates will be. In fact, large errors in the data sometimes make it impossible to accurately estimate any parameters. The problems associated with data noise are partially circumvented by introducing weighting. Weighting is simply a way to indicate the relative importance of each data point to the total SSQ. Weighted SSQ are calculated as follows:

$$SSQ = \sum_{i=1}^{n} W_i (Y_i - \hat{Y}_i)^2$$

Where Wi indicates the weight for the given data point. Typically, inverse variance is used to weight the data. Inverse variance assumes that the user knows something about the true error in the data. For example, it is standard practice to know the coefficient of variation in an assay. This number could be used as a known weighting factor. The important thing to remember is that it is always assumed that the noise is random with a normal distribution. Bias will be introduced to the extent that this assumption is violated.

If you have no idea about the actual variance in the data, an empirical approach may be used. One such approach is to smooth the data with some smoothing function, and estimate the variance by comparing the measured and smoothed data. Smoothing can be done using a moving average, polynomial, spline function, or optimal segments.