Paper SA-10-2011

PROC NLMIXED for Basic Non-Linear Regression

Keith Dunnigan
Strategic Staffing Solutions, St Louis, MO

Abstract

PROC NLMIXED is a SAS procedure which can be used to analyze nonlinear regression models containing more than one variance component. This presentation provides an introductory primer to PROC NLMIXED regarding its basic use and syntax in the context of a basic non-linear regression model. The syntax for a non-linear regression analysis will be applied to example data from a hypothetical dose-response study.

Introduction:

The purpose of this presentation is to provide an introduction into the required syntax of PROC NLMIXED needed to analyze a basic non-linear regression model. After providing a brief description of the procedure, the syntax for a basic design and related outputs will be presented.

Background – PROC NLMIXED

PROC NLMIXED is a SAS procedure which can be used to analyze nonlinear regression models that may contain random effects (or variance components other than the residual). The model analyzed may involve a mixture of fixed and random parameters, any of which may enter the model either linearly or nonlinearly. The residual may be modeled as either a normal, binomial or poisson random variable. Other random effects in the model may be modeled only by the normal distribution.

In the context of other SAS procedures, it is most similar to PROC NLIN and differs from it mainly in its ability to accommodate random effects. If no variance components exist in the model other than the residual, PROC NLIN may be a more appropriate procedure. PROC NLMIXED features a random statement which is similar to that used in PROC MIXED. Unlike PROC MIXED, it has no class statement and does not produce traditional ANOVA analyses or outputs such as LSMEANS.

The basic syntax used in PROC NLMIXED to perform a nonlinear regression will now be described in the context of a worked out example involving data from the following hypothetical dose response study.

Dose Response Study

A clinical trial on an experimental drug is conducted to study the effect of dosage on the percentage change from baseline (%CFB) concentration of a certain analyte in the bloodstream. Subjects entering this trial are randomized to one of 3 doses of an experimental drug. Subjects have their blood drawn before dosing and at 3 selected time-points after dosing has occurred. The concentration of the analyte is measured at each of these time-points. This dataset is illustrated here for several subjects:

Dose Response Example Data

SUBJECT_ID	TIME	dose	BASELINE	Pct_Chg_base
1	1	75	50.6789	81.7286
	5	75	50.6789	14.5232
	10	75	50.6789	41.9613
2	1	25	50.9555	68.8034
	5	25	50.9555	78.4140
	10	25	50.9555	85.7684
3	1	0	41.0876	40.0383
	5	0	41.0876	21.9002
	10	0	41.0876	38.6676

This data is analyzed using the following model:

$$Y_{ij} = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] * \left[1 - e^{(\beta_4 * t)}\right] + \gamma_i + e_{ij} \qquad (1)$$

$$\gamma_i \sim N(0, \sigma_s^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

Where Y is the percent change from baseline in the drug concentration measured on subject i at time-point j. There are i*j measurements, each having dosage d, time t and baseline concentration BL. The model contains fixed parameters β_0, β_1, β_2, β_3, β_4, random effects γ_i, and ε_{ij} (subject, residual), and variables BL, d, and t which are treated as covariates (baseline, dosage, time). The random effects are both distributed as standard normal random variables with variance parameters σ_s^2 and σ_e^2 respectively. β_0, γ_i, and ε_{ij} enter the model linearly, all other effects and parameters enter nonlinearly. The basic syntax required to fit this model and estimate the parameters will now be presented.

SAS Example 1 – Parameter Estimates for Dose Response Study

The syntax for analyzing this model is shown below.

```
proc nlmixed data = conc_data;
  term3 = (beta2*dose)/(beta3+dose);
  term4 = 1 - exp(beta4 * time);
 parms beta0=3.14
       beta1 = -0.015
       beta2=319
       beta3=57.3634
       beta4 = -0.904
       s2s = 385.7
       s2e = 385.7;

  model pct_chg_base ~ normal (beta0 + ( ( exp(beta1 * baseline) ) *
                        term3 * term4) + subj, s2e);
  random subj ~ normal(0,s2s) subject=subject_id;
run;
```

This syntax consists of the parms, model, and random statements as well as several programming statements, which are described below:

1) Parms Statement: The parms statement defines the parameters in the model. It also specifies the initial values (or initial estimates) for each parameter, which SAS uses to begin the estimation algorithm. SAS calculates the final estimates for each parameter at the conclusion of the algorithm.

2) Programming statements: PROC NLMIXED allows programming statements within the procedure. These may be used to define variables, which can be subsequently used within other procedure statements in order to simplify them. These statements are similar to those used within data-steps and any valid SAS statement is allowed involving parameters, covariates, and fixed constants. (The variables created exist only within the NLMIXED procedure). In our example the 'term3', 'term4', and 'pred' variables are defined using programming statements and subsequently used in the model statement.

3) Model Statement: The model statement specifies the conditional distribution of the data given the random effects. It involves the dependent variable from the input data set, a tilde (~), and then a distribution stated with its parameters listed in parenthesis and separated by commas: Y ~ Distr (par1, par2, …). Several distributions are possible within PROC NLMIXED. In our example, we use the normal distribution which has two parameters (μ,σ^2) and the model statement takes on the form Y ~ Normal (mean, variance):

    a. First the distribution is stated. We model a normal distribution.

    b. Next the mean of the model in equation 1, conditional on the random effect "subj", is specified within the parentheses, and before the comma. (Note that since the mean of e_{ij} is 0, the error term disappears in this expression, while the random term 'subj' remains since this is a conditional distribution). This expression involves parameters, covariates

and fixed constants as well as the variables defined within the procedure ('term3', 'term4', and 'pred').

c. Finally the variance of the model distribution is specified, following the comma. In examing equation 1, it can be seen that the variance of the model is that of the residual, since the conditioned random effect 'subj' is held constant. This residual variance is specified here as s2e.

4) Random Statement: The random statement is used to specify the random effects included in the model statement and to identify their distributions. It specifies the distribution in a similar way as does the model statement. Here the subject effect is specified as a normal distribution with mean 0 and variance parameter equal to s2s.

The "subject = option" functions similarly to the "subject = option" within the random and repeated statements of PROC MIXED. It specifies the units within the data which will define the values of the random effect. Here these values correspond to each subject within the data. Other items to note:

i. Sort Order: PROC NLMIXED assumes that a new value of the random effect occurs whenever the unit of the SUBJECT= variable changes from the previous observation. As an example, if 6 records with subject values (1,2) are arranged in the data in subject order 1,2,1,2,1,2, SAS will identify this as 6 different subject values for purposes of estimating the random effect. *The input data set should always be sorted according to the 'subject=' variable prior to running PROC NLMIXED.*

ii. More than one random effect may be specified.

iii. Only the normal distribution is available for modeling random effects.

Selected output from this procedure is shown below and details the final parameter estimates.

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
beta0	3.6476	4.1231	114	0.88	0.3782	0.05	-4.5202	11.8155	0.00174
beta1	-0.00967	0.003481	114	-2.78	0.0064	0.05	-0.01656	-0.00277	5.657628
beta2	297.87	124.45	114	2.39	0.0183	0.05	51.3291	544.41	0.011116
beta3	75.1038	46.3144	114	1.62	0.1076	0.05	-16.6447	166.85	0.002028
beta4	-3.1775	0.8719	114	-3.64	0.0004	0.05	-4.9047	-1.4503	0.006332
s2s	414.55	64.0760	114	6.47	<.0001	0.05	287.62	541.49	-0.01609
s2e	298.07	29.3731	114	10.15	<.0001	0.05	239.88	356.26	-0.00723

Dose Response Study – Predictions and Estimates

Once the model described in equation 1 is fit to the dose response data, several output measures are desired, which may be found using PROC NLMIXED. From the model equation, it can be seen that the dependent variable is a function of baseline, dosage, and time, with a subject effect. The dependent variable (concentration CFB) may be estimated from the model given all 4 values (equation 2). We may also make estimates across all subjects for any given values of BL, d, and t by dropping the random subject effect as shown in equation 3.

$$\hat{Y}_i(BL, d, t) = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] * \left[1 - e^{(\beta_4 * t)}\right] + \gamma_i \qquad (2)$$

$$\hat{Y}(BL, d, t) = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] * \left[1 - e^{(\beta_4 * t)}\right] \qquad (3)$$

In our study we would like to make overall estimates of the CFB concentration for each dose. We can do this by applying equation 3 and substituting in the mean values of the time and baseline variables (over the input data):

$$\hat{Y}(d \mid \overline{BL}, \bar{t}) = \beta_0 + e^{(\beta_1 * \overline{BL})} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] * \left[1 - e^{(\beta_4 * \bar{t})}\right] \qquad (4)$$

Finally we would like to measure the effect of changing the dosage on the CFB concentration. We do this by estimating the difference in the two estimates for two given doses:

$$\left(\hat{Y}_i(d_1 \mid \overline{BL}, \bar{t}) - \hat{Y}_i(d_2 \mid \overline{BL}, \bar{t}) = e^{(\beta_1 * \overline{BL})} * \left[1 - e^{(\beta_4 * \bar{t})}\right] * \left\{\left[\frac{\beta_2 * d_1}{\beta_3 + d_1}\right] - \left[\frac{\beta_2 * d_2}{\beta_3 + d_2}\right]\right\} \qquad (5)$$

We can make estimates of the type shown in equation 2 in PROC NLMIXED only by using the PREDICT statement. We can make estimates of the other three types using either PREDICT or ESTIMATE statements, although it is easier to do so using the ESTIMATE statement. It is useful to note that the PROC REG procedure (used in linear regression) does not have the equivalent of an estimate statement in its syntax.

SAS Example 2 – Predictions from the Dose Response Study

The PREDICT statement allows one to construct predictions of an expression for each observation in the input data set. Any valid SAS programming expression involving the input data set variables, parameters, and random effects is allowed. Predicted values are computed using the parameter estimates and empirical Bayes estimates of the random effects. Standard errors of prediction are computed using the delta method (Billingsley 1986, Cox 1998). Results are placed in an output data set using the 'out=' option.

We start with the basic syntax of SAS Example 1 and add the following two statements to make predictions of the type described in equation 2. Note that these contain the random

subject effect (highlighted). A segment of the resulting SAS output dataset ('predout') is shown for subject 15, time 5, and has covariates: baseline = 83.7685 and dose = 25.

```
pred = beta0 + ( ( exp(beta1 * baseline) ) * term3 * term4) + subj;
predict pred out = predout;
```

SUBJECT_ ID	TIME	dose	BASELINE	Lower Confidence Limit	Predicted Value	Upper Confidence Limit	Pr > |t|	Pct_Chg_ base
15	5	25	83.4542	33.0971	51.0177	68.9384	.000000126	52.6158

If we omit the subject effect in the above syntax, we can make predictions across all subjects (equation 3). The following output results, and illustrates the substantial impact of subject to subject variability in this data.

SUBJECT_ ID	TIME	dose	BASELINE	Lower Confidence Limit	Predicted Value	Upper Confidence Limit	Pr > |t|	Pct_Chg_ base
15	5	25	83.4542	25.2508	36.8455	47.9402	1.5006E-9	52.6158

SAS Example 3 – Estimates from the Dose Response Study

The ESTIMATE statement may be used to compute an estimate that is a function of the parameter values. Like the predict statement it enables one to estimate SAS expressions involving parameters, constants, and programmed variables. Unlike the predict statement it does not allow random effects to be included in these expressions. PROC NLMIXED also computes approximate standard errors for these estimates using the delta method.

The following code lines are added to the syntax of SAS Example 1 to calculate estimates at each dose level (equation 4), and for differences between the doses (equation 5):

```
term4b = (beta2*0)/(beta3+0);
term5 = (beta2*25)/(beta3+25);
term6 = (beta2*75)/(beta3+75);

term8 = 1 - exp(beta4 * 5.34);
term9 = 1 - exp(beta4 * 5);

/*estimate response at mean baseline and different dose levels*/

estimate "Dose 0" beta0   + ( ( exp(beta1 * 42.5) ) *
                             ((beta2*0)/(beta3+0)) * term8);
estimate "Dose 25" beta0  + ( ( exp(beta1 * 42.5) ) *
                             ((beta2*25)/(beta3+25)) * term8);
estimate "Dose 75" beta0 + ( ( exp(beta1 * 42.5) ) *
                            ((beta2*75)/(beta3+75)) * term8);
```

```
/* Estimate differences between dose levels at mean baseline */

  estimate 'Dose25 - Placebo'  (exp(beta1 * 42.5) ) * term8 *
                               (term5 - term4b);
  estimate 'Dose75 - Placebo' (exp(beta1 * 42.5) ) * term8 *
                              (term6 - term4b);

  estimate "D 25,T 5,BL 83.7685" beta0   + ( ( exp(beta1 * 83.7685) ) *
                                   ((beta2*25)/(beta3+25)) * term9);
```

The output produced from these statements is shown below. If we compare the last line (highlighted) with the same prediction made using the estimate statement (excluding the random effect), we can see the results are identical. This illustrates the fact that the PREDICT and ESTIMATE statements calculate the prediction limits using similar methods.

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Dose 0	3.6476	4.1231	114	0.88	0.3782	0.05	-4.5202	11.8155
Dose 25	52.9728	5.4131	114	9.79	<.0001	0.05	42.2496	63.6961
Dose 75	102.33	6.3459	114	16.13	<.0001	0.05	89.7609	114.90
Dose25 - Placebo	49.3252	7.6882	114	6.42	<.0001	0.05	34.0949	64.5555
Dose75 - Placebo	98.6845	6.6942	114	14.74	<.0001	0.05	85.4234	111.95
D 25,T 5,BL 83.7685	36.7448	5.6092	114	6.55	<.0001	0.05	25.6330	47.8565

Further Applications of the Estimate Statement

It is useful to note that the SAS Procedures principally used for linear and non-linear regression (REG, GLM, NLIN) do not have an equivalent estimate statement to NLMIXED. Neither REG nor NLIN include an estimate statement. GLM does have an estimate statement, but it estimates only *linear* functions of the model parameters.

Normally we should fit linear regression models with REG or GLM, and simpler nonlinear models (without multiple variance components) using NLIN. However there may be circumstances where we wish to fit a simpler model (such as a linear regression), and yet still estimate a nonlinear function of the resulting model parameters. In those situations we may still consider using NLMIXED. We would first fit the simpler model, then use the estimates obtained as initial values into NLMIXED.

Pharmacokinetic Dose Proportionality Example:

A pharmacokinetic (PK) parameter is considered dose proportional if this quantity doubles in the bloodstream each time the dosage doubles. In other words, it follows the relationship $PK = \beta_0 + \beta_1 Dose$, where the intercept β_0 is zero. Traditionally dose proportionality is tested by fitting a simple linear regression model and testing the slope and intercept parameters against zero.

Another method of testing dose proportionality involves constructing confidence intervals around the ratio formed by β_0 / β_1 and then comparing the endpoints generated

with an equivalence region (Smith 2000). Methods of estimating β_0/β_1 have involved either Fieller's Theorem or Bootstrapping techniques (Smith 2000, Dunnigan 2005). The ESTIMATE statement of PROC NLMIXED could provide yet another method method of estimation.

Finding Initial Values

Unlike linear regression models, Non-linear models do not have closed form solutions and require iterative computational methods. It is necessary to specify initial values (IVs) for the parameters. Specifying incorrect IVs can cause the algorithm not to converge, or worse, to converge to an incorrect point such as a local (as opposed to global) minimum . Therefore it is necessary to exercise great care when choosing initial values for PROC NLMIXED and we should choose IV's as close as possible to the final values. This becomes even more critical as models become more complex, involving many parameters.

Finding the initial values is the most difficult and time consuming part of analyzing a non-linear regression model. There are many approaches to finding the IVs and no particular 'right' or 'wrong' methods. Our purpose here is simply to describe the approach used with this particular model and dataset. This process was used to find the initial values given in the syntax to SAS Example 1.

The first step is to examine the model and perform descriptive analyses.

$$\hat{Y}(BL,d,t) = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] * \left[1 - e^{(\beta_4 * t)}\right]$$

Estimating Beta0:

We can see from the model above that when the dose is zero, the 3 multiplicative terms drop out, leaving only Beta0. We then compute an initial value (IV) of Beta0 by averaging the dependent variable over the dataset for all records where the dose is zero.

$$\widehat{\beta_0} = \{\hat{Y} \,|d = 0\}$$

Estimating Beta2, Beta3

We can see that the model is a function of baseline, dose, and time: Y = f(BL,d,t)). We start by examining scatterplots of Y*BL, Y*d, and Y*t. In examining these we observed a significant relationship between Y and dose, but little between Y and BL (a slightly decreasing line), and even less between Y and time (an almost flat line).

Since the right side of the model statement consists of the product of 3 functions, we postulate that the BL and t multiplicative terms contribute little and we can substitute in one (1) for each expression. We also substitute in Ybar and dbar (the mean values taken over all the data). The resulting equation is given by:

$$\bar{Y} = \widehat{\beta_0} + \left[\frac{\beta_2 * \bar{d}}{\beta_3 + \bar{d}}\right]$$

It is difficult to postulate values for β_2 and β_3 separately, since it is their ratio that is really driving this equation. We assume that they should be different from one another.., but how different? It is safest is to assume they are equal and let the modeling process tell us how unequal they are. Therefore we set $\widehat{\beta_2} = \widehat{\beta_3} = \hat{\beta}$ and solve for $\hat{\beta}$:

$$\bar{Y} = \widehat{\beta_0} + \left[\frac{\hat{\beta} * d}{\hat{\beta} + d}\right], \quad \widehat{\beta_2} = \widehat{\beta_3} = \hat{\beta} = \frac{(\bar{Y} - \widehat{\beta_0})\bar{d}}{\bar{d} - (\bar{Y} - \widehat{\beta_0})}$$

Re-Estimating Beta2, Beta3:

With these initial values we will fit model A given below:

$$Y = \beta_0 + \left[\frac{\beta_2 * d}{\beta_3 + d}\right] \quad \text{(A)}$$

We seek out the new IVs recursively by:

1) Setting $\widehat{\beta_0}$ and $\widehat{\beta_2}$ as before and grid sampling over $\widehat{\beta_3}$. We begin with a wide grid and search for the region with the lowest SSE. We then repeat this procedure again, narrowing in on the region(s) of interest. We repeat this process iteratively until we have identified the correct starting value, which occurs at the operating point with the lowest SSE.

2) Set $\widehat{\beta_0}$ as before, set $\widehat{\beta_3}$ to its new value, and grid sample over $\widehat{\beta_2}$ iteratively.

3) At the conclusion of this analysis, we run model A, setting $\widehat{\beta_0}$ as before and setting $\widehat{\beta_2}, \widehat{\beta_3}$ to the new values obtained. The final model parameters estimated form the new values of $\widehat{\beta_0}, \widehat{\beta_2}$, and $\widehat{\beta_3}$ that we use in the next step.

The syntax and SAS output below illustrates the analyses described in 1 above, and the grid sampling procedures in general used for determining IVs. In the output shown, the regions highlighted should be expanded on further in repeated analyses.

```
parms beta0=1.5394 beta2=110.941 beta3=-5000 to 5000 by 100;
model pct_chg_base = beta0 + term3;
```

Grid Search

beta2	beta3	beta0	Sum of Squares
110.9	-1300.0	1.5394	7422298
110.9	-1200.0	1.5394	7977211
110.9	-1100.0	1.5394	8755009
110.9	-1000.0	1.5394	9917711
110.9	-900.0	1.5394	11825520
110.9	-800.0	1.5394	15448638
110.9	-700.0	1.5394	24383902
110.9	-600.0	1.5394	64897970
110.9	-500.0	1.5394	2700624
110.9	-400.0	1.5394	28505414
110.9	-300.0	1.5394	7338953
110.9	-200.0	1.5394	7448457
110.9	-100.0	1.5394	1871954
110.9	0	1.5394	1279156
110.9	100.0	1.5394	1045285
110.9	200.0	1.5394	1493361
110.9	300.0	1.5394	1817141
110.9	400.0	1.5394	2063664

Estimating Beta1

Fit model B shown below (which adds the BL term to model A). Set $\widehat{\beta_0}$, $\widehat{\beta_2}$, and $\widehat{\beta_3}$ at the values obtained from the final output of model A. Grid search over $\widehat{\beta_1}$ iteratively. Once the initial value is identified, run model B again, using the final estimated parameters for beta0, beta1, beta2, and beta3 as initial values in the next step.

$$Y = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right] \qquad \text{(B)}$$

Estimating Beta4

Fit model C shown below, which adds the term t. Set $\widehat{\beta_1}$, $\widehat{\beta_0}$, $\widehat{\beta_2}$, and $\widehat{\beta_3}$ at the final parameter estimate values obtained from model B above. Grid search over $\widehat{\beta_4}$ iteratively. Once the initial value is identified, run model C again, using the final estimated parameters for beta0, beta1, beta2, beta3, and beta4 as initial values into the next step.

$$Y = \beta_0 + e^{(\beta_1 * BL)} * \left[\frac{\beta_2 * d}{\beta_3 + d}\right]\left[1 - e^{(\beta_4 * t)}\right] \qquad \text{(C)}$$

Initial Values for Variance Components

The initial values for all the fixed parameters have been determined. Fit model C using PROC NLIN and record the output MSE. Estimate the IV's for both variance components (subject and residual) as ½ the MSE:

$$\widehat{\sigma_s^2} = \widehat{\sigma_e^2} = MSE/2$$

At this point all initial values have been found and are ready for inclusion in the final NLMIXED analysis, as shown in SAS Example 1.

Summary

In this presentation we have illustrated the basic syntax of PROC NLMIXED for non-linear regression by presenting the analysis of example data from a hypothetical dose-response study. The MODEL, PARMS, RANDOM, ESTIMATE, PREDICT and programming statements were described and applied through this example. Example syntax and output were shown.

As we have described, PROC NLMIXED is a useful procedure for performing non-linear regression analyses. It is similar to Proc NLIN, and yet also allows us to fit models containing more than one variance component. It also expands the number of probability distributions available for modeling the residual. We may wish to choose NLMIXED over NLIN for fitting models involving a non-normal residual and/or for those including random effects.

The procedure includes useful statements (ESTIMATE, PREDICT) for estimating linear or nonlinear functions of model parameters, covariates, constants, and even random subject effects (PREDICT). This flexibility may make NLMIXED useful in some cases, even with simpler statistical models. Finally it includes a RANDOM statement with very similar syntax to that found in PROC MIXED, and also requires sorting the data by the 'subject =' variable prior to running the code.

References

1) The NONLIN Procedure. SAS STAT ® 9.22 Users Guide . (06/22/2010). http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#nlmixed_toc.htm

2) Billingsley, P. (1986), *Probability and Measure*, Second Edition, New York: John Wiley & Sons, Inc.

3) Cox, C. (1998), "Delta Method," *Encyclopedia of Biostatistics*, Eds. Peter Armitage and Theodore Colton, New York: John Wiley, 1125–1127.

4) Smith BP, Vandenhende FR, DeSante KA, Farid NA, Welch PA, Callaghan JT, and Forgue ST. Confidence Interval Criteria for Assessment of Dose Proportionality. Pharm Res 2000; 17:1278-1283.

5) Dunnigan, Keith (2005) Using SAS to Assess Dose Proportionality in Dose Escalation Studies. 2005 MWSUG Conference Proceedings.

Contact Information

The author may be contacted at:

Keith Dunnigan
Dunnigan_k@yahoo.com