Annotated SAS Code for Contrasts

To compute and test for significance of a contrast in SAS, we add a "contrast" statement to the GLM procedure. We give some examples of this below.

Example 1. This is the example on the yield of beans discussed in Part 6 of the Analysis of Variance module. Beans were planted in 3 different densities: 10 plants per plot, 20 plants per plot, and 30 plants per plot. Yield was measured on each plot. We want to test if the linear trend or if the quadratic trend is significant.

In SAS, the trends are tested with contrast statements.

```
data beans;
input Density Yield @@;
datalines;
10 11.9 10 11.1 10 13.9
20 16.1 20 15.0 20 16.9
30 20.3 30 18.6 30 17.5
;

ods graphics on;
proc glm data=beans;
 class Density;
 model Yield = Density;
 contrast 'Linear' Density -1 0 1;
 contrast 'Quadratic' Density 1 -2 1;
 run;
ods graphics off;
quit;
```

Let's consider the components of the linear contrast statement.

```
contrast 'Linear' Density -1 0 1;
```

`contrast` asks that a contrast be computed.

`'Linear'` gives a name to the contrast. It must be in quotes and it can be whatever we want it to be. We could call the contrast `'George'` if we wanted to.

`Density` is the name of the quantitative predictor variable. It must be the same as a term on the right side of the model statement.

`-1 0 1` are the coefficients of the contrast. In this case, these are the coefficients for computing the linear contrast involving 3 equally spaced treatments. Don't forget the semi-colon at the end of the statement.

The quadratic contrast is constructed in a similar way.

Edited Output for Example 1

In addition to the usual Proc GLM output, there are two additional lines (one for each contrast statement).

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Linear	1	63.37500000	63.37500000	38.18	0.0008
Quadratic	1	0.40500000	0.40500000	0.24	0.6389

When we examined these contrasts "by hand" (in Part 6 of the Analysis of Variance module), we calculated test statistics based on the t distribution, while the ones given here are based on the F distribution. These results are consistent because the square of t is the same as F. For example, the linear contrast we calculated t = 6.5 and the result given here is F = 38.18 = 6.5^2 (so F = t^2). From the t distribution, we concluded that there is a linear trend. For the F test, the p-value is 0.0008 so we reject the hypothesis that the contrast equals 0, i.e. we reject the hypothesis that there is not a linear trend. Thus both the t test and F test result in the same conclusion *because it is the same test*.

Contrasts are not limited to linear and quadratic trends. They can be used to test more general hypotheses of treatment means.

Example 2. Suppose we wish to compare the mean yield of density 10 to the average of the yields of densities 20 and 30. The contrast is $\mu_1 - (\mu_2 + \mu_3)/2$ when expresses in terms of population means. Thus the contrast coefficients are 1, –0.5, and –0.5. The following contrast statement would test for significance of the contrast

```
contrast '1 vs avg of 2 and 3' Density 1 -0.5 -0.5;
```

We could also multiply the coefficients by 2 to convert to whole numbers. Thus we could also write the contrast statement as

```
contrast '1 vs avg of 2 and 3' Density 2 -1 -1;
```

Example 3: Caution. To use the contrast statement, the sum of the coefficients must be 0. There is no allowable error for rounding. Suppose we have 4 treatments and wish to compare treatment 1 to the average of treatments 2, 3, and 4. This would give us the following contrast for population means: $\mu_1 - (\mu_2 + \mu_3 + \mu_4)/3$. However, if we were to set the coefficients as 1, -.333, -.333, and -.333 we would get an error because they do not add exactly to 0. We could use 1, -.333, -.333, and -.334 (changing the last -.333 to -.334), but a better way is to multiply the coefficients by 3 and use 3, -1, -1 and -1.

Example 4: Caution. Suppose we have three pain medications labeled Control, Aspirin and Tylenol. Suppose we wish to compare control response to the average response of Aspirin and Tylenol. SAS will put the treatments in alphabetic order: Aspirin, Control, Tylenol. Thus the coefficients for Control vs. the average of Aspirin and Tylenol should be -.5, 1, -.5 (or -1, 2, -1). If we were to use 1, -.5, -.5 this would compare Aspirin to the average of Control and Tylenol. If in doubt as to the order SAS is using, look at the output of the LSMEANS statement. This is the order of the treatments you should use in the contrast statement.

Coefficients for linear and quadratic trends for up to 8 equally spaced treatments

Number of Treatments	Contrast	Coefficients							
3	linear	-1	0	1					
	quadratic	1	-2	1					
4	linear	-3	-1	1	3				
	quadratic	1	-1	-1	1				
5	linear	-2	-1	0	1	2			
	quadratic	2	-1	-2	-1	2			
6	linear	5	-3	-1	1	3	5		
	quadratic	5	-1	-4	-4	-1	5		
7	linear	-3	-2	-1	0	1	2	3	
	quadratic	5	0	-3	-4	-3	0	5	
8	linear	-7	-5	-3	-1	1	3	5	7
	quadratic	7	1	-3	-5	-5	-3	1	7

When the treatments are not evenly spaced, the coefficients for linear and quadratic trends depend on the levels of the treatments and their sample sizes. These are not calculated by hand -- a SAS program is given below. In the program, 'a' contains the levels of the treatments and 'b' contains the sample sizes. 'orpol' is a built-in SAS function to calculate orthogonal polynomials.

For example, the following code generates the coefficients for linear and quadratic trends for data in which the quantitative treatments are 10, 20, 40 and 80 and the sample sizes are 5 for each treatment.

```
proc iml;
a = {10 20 40 80};
b = {5 5 5 5};
coeff = orpol(a,2,b);
print a;
print b;
print coeff;
run;
quit;
```

a			
10	20	40	80

b			
5	5	5	5

coeff		
0.2236068	-0.229366	0.2368565
0.2236068	-0.14596	-0.047371
0.2236068	0.0208514	-0.343442
0.2236068	0.3544745	0.1539567

The first column of the coefficients are constants that can be ignored. The second column has the coefficients for the linear trend (-0.2293, -0.14596, ...) . The third column has the coefficients for the quadratic trend (0.2368, -0.0473, ...). Round the numbers so they add to 0.