PROC GLM AND PROC MIXED CODES FOR TREND ANALYSES FOR ROW-COLUMN DESIGNED EXPERIMENTS

BU-1491-M **June, 2000**

Walter T. Federer
Dept. of Biometrics
Cornell University
Ithaca, NY 14853
wtf1@cornell.edu

and

Russell D. Wolfinger
SAS Institute, Inc.
R52, SAS Campus Drive
Cary, NC 27513

Keywords: trend analyses, exploratory model selection, differential gradients, orthogonal polynomial regression.

Abstract

A SAS code is written for five different response models for a row-column laid out experiment. These are useful in exploratory model selection to determine which model best fits the spatial variation present in the experiment. The five models are for a randomized complete block, a r-w-column, differential gradients within rows (columns), orthogonal polynomial regression of row and column order and interactions, and a mixture of row-column and regression interactions.

Title: PROC GLM AND PROC MIXED CODES FOR TREND ANALYSES FOR ROW-COLUMN DESIGNED EXPERIMENTS

Authors: W. T. Federer, 434 Warren Hall, Cornell University, Ithaca, NY 14853, e-mail WTF1@cornell.edu and R. D. Wolfinger, SAS Institute, Inc. R52, SAS Campus Drive, Cary, NC 27513.

Purpose: This program may be used for a variety of response models for a row-column laid out experiment. The example used to illustrate the steps in the program is for a randomized complete block design (RCBD) which was laid out as an eight row by seven column field experiment. The experiment with data is described in Federer, W. T. and C. S. Schlottfeldt (1954), Biometrics 10:282-290. The data are totals of 20 plant heights in centimeters for seven different treatments. Since the experiment was laid out an eight row by seven column arrangement, an RCBD analysis may not be appropriate. The SAS code is written to compare five different response models for accounting for the spatial variation present. There appeared to be variation oriented differently than the row-column layout. SAS PROC GLM and PROC MIXED codes are presented for standard textbook analyses of variance for a RCBD and for a row-column design. These are followed by codes for trend analyses using standardized orthogonal polynomial regressions for rows and columns and for interaction of row and column regressions. A trend model using row, column, and interactions of row and column regressions appears to control the variation for this experiment. A PROC GLM analysis of variance and residuals is useful in exploratory model selection of a model that takes account of the spatial variation in the experiment. Then, a PROC MIXED analysis is used to recover information from the random effects.

References: Federer, W. T. (1998). Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. Biometrics 54(2):471-481.
Federer, W. T. and R. D. Wollfinger (1998). SAS PROC GLM and PROC MIXED code for recovering inter-effect information. Agronomy Journal 90:545-551.

SAS Code:

```
/*---input the data---*/
data colrow;
  input height row col trt;
  /*---rescale data for stability---*/
  y = height/1000;
  datalines;
1299.2  1 1 6
875.9   1 2 7
 960.7  1 3 4
1004.0  1 4 3
1173.2  1 5 1
1031.9  1 6 2
1421.1  1 7 5
1369.2  2 1 2
 844.2  2 2 5
 968.7  2 3 6
 975.5  2 4 7
1322.4  2 5 3
1172.6  2 6 1
1418.9  2 7 4
1169.5  3 1 1
 975.8  3 2 5
 873.4  3 3 3
 797.8  3 4 7
1069.7  3 5 2
1093.3  3 6 6
1169.6  3 7 4
1219.1  4 1 6
 971.7  4 2 1
```

```
607.6   4 3 7
1000.0  4 4 4
1343.3  4 5 2
999.4   4 6 5
1181.3  4 7 3
1120.0  5 1 6
827.0   5 2 7
671.9   5 3 4
972.2   5 4 3
1083.7  5 5 1
1146.9  5 6 2
993.8   5 7 5
1031.5  6 1 7
846.5   6 2 2
667.8   6 3 4
853.6   6 4 3
1087.1  6 5 1
990.2   6 6 5
1021.9  6 7 6
1076.4  7 1 2
917.9   7 2 1
627.6   7 3 5
776.4   7 4 6
960.4   7 5 3
852.4   7 6 7
1006.2  7 7 4
1099.6  8 1 4
947.4   8 2 5
787.1   8 3 2
898.3   8 4 1
1174.9  8 5 3
1003.3  8 6 6
947.6   8 7 7
run;

/*---code to construct orthogonal polynomials---*/
proc iml;
  /*---7 columns and up to 6th degree polynomials---*/
  opn4=orpol(1:7,6);
  opn4[,1] = (1:7)`;
  op4= opn4;
  create opn4 from opn4[colname={'col' 'c1' 'c2' 'c3' 'c4' 'c5'
    'c6'}];
  append from opn4;
  close opn4;
  /*---8 rows and up to 7th degree polynomials---*/
  opn3=orpol(1:8,7);
  opn3[,1] = (1:8)`;
  op3 = opn3;
  create opn3 from opn3[colname={'row' 'r1' 'r2' 'r3' 'r4' 'r5'
    'r6' 'r7'}];
  append from opn3;
  close opn3;
run;
/*---merge in polynomial coefficients---*/
data rcbig;
```

```
  set colrow;
  idx = _n_;
proc sort data=rcbig;
  by col;
data rcbig;
  merge rcbig opn4;
  by col;
proc sort data=rcbig;
  by row;
data rcbig;
  merge rcbig opn3;
  by row;
proc sort data = rcbig;
  by idx;
run;
/*---3d plot of data, one can also substitue row and column variables as well as residuals for
   y to see how they model the trend---*/
proc g3d data=rcbig;
  plot row*col=y / rotate=20;
run;
/*---standard rcbd analysis with rows as blocks; treatments are
   not significantly different---*/
/*---fixed-effects row model for RCBD---*/
proc glm data=rcbig;
  class row col trt;
  model y = row trt;
  output out=subres r=resid;
run;
/*---standard row-column analysis fits much better than RBCD, and
   now treatment 7 is significantly different---*/
/*---fixed-effects row-column model---*/
proc glm data=rcbig;
  class row col trt;
  model y = row col trt;
  output out=subres r=resid;
run;
/*---model for random differential gradients within rows; does not fit
   as well as row-column model, but results are similar---*/
/*---fixed-effects model for gradients within rows ---*/
proc glm data=rcbig;
  class row col trt;
  model y = trt row c2*row c3*row c4*row;
  output out=subres r=resid;
run;
/*---Fixed-effects polynomial model; it may be that a trend and analysis
   is desired in that only certain polynomial regressions are needed to
   explain the row and column variation.  Also, since spatial variation
   may not be in the row-column orientation of the experiment,
   interactions of regressions may be needed to account for this type
   of spatial variation.  Of the 13 polynomial regressions for rows and
   columns and the 16 interactions ci*rj, for i, j = 1, 2, 3, and 4,
   those that had F-values greater than F at the 25% level were
   retained in the response model.---*/
proc glm data=rcbig;
  class row col trt;
  model y = trt c1 c2 c3 c5 r1 r2 r3 r5 r6 r7 c1*r1 c2*r1 c2*r3 c3*r2 c4*r1 c4*r2;
```

```
  output out=subres r=resid;
run;
/*---spatial covariance model---*/
proc mixed data=rcbig;
  class row col trt;
  model y = trt / ddfm=res;
  random c1 c2 c3 c5 r1 r2 r3 r5 r6 r7 c1*r1 c2*r1 c2*r3 c3*r2 c4*r1 c4*r2;
  lsmeans trt / diff adjust=tukey;
run;
/*---Since the row and column variations were quite un-patterned, i.e.,
   only c4, c6, and r4 were not in the model, the following analysis
   may be more appropriate for this data set.---*/
proc glm data=rcbig;
  class row col trt;
  model y = row col trt c1*r1 c2*r1 c2*r3 c3*r2 c4*r1 c4*r2;
run;
/*---spatial covariance model---*/
proc mixed data=rcbig;
  class row col trt;
  model y = trt / ddfm=res;
  random row col  c1*r1 c2*r1 c2*r3 c3*r2 c4*r1 c4*r2
  repeated / type=sp(exp)(row col) subject=intercept;
  lsmeans trt / diff adjust=tukey;
run;
```

An abbreviated output from this code is presented below:

RCBD ANOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	13	0.66219035	0.05093772	1.69	0.1004
Error	42	1.26958627	0.03022824		
Corrected Total	55	1.93177662			

R-Square	C.V.	Root MSE	Y Mean
0.342788	17.17205	0.173863	1.012475

Dependent Variable: Y

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	1.84	0.1056
TRT	6	0.27387545	0.04564591	1.51	0.1985

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	1.84	0.1056
TRT	6	0.27387545	0.04564591	1.51	0.1985

Row-column ANOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	1.66711058	0.08774266	11.93	0.0001
Error	36	0.26466604	0.00735183		
Corrected Total	55	1.93177662			

R-Square	C.V.	Root MSE	Y Mean
0.862993	8.468638	0.085743	1.012475

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	7.55	0.0001

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
COL	6	1.15907213	0.19317869	26.28	0.0001
TRT	6	0.11972355	0.01995392	2.71	0.0281
Source	DF	Type III SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	7.55	0.0001
COL	6	1.00492023	0.16748671	22.78	0.0001
TRT	6	0.11972355	0.01995392	2.71	0.0281

Gradients within rows ANOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	37	1.72819875	0.04670807	4.13	0.0011
Error	18	0.20357788	0.01130988		
Corrected Total	55	1.93177662			

R-Square	C.V.	Root MSE	Y Mean
0.894616	10.50376	0.106348	1.012475

Dependent Variable: Y

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	6	0.27387545	0.04564591	4.04	0.0098
ROW	7	0.38831490	0.05547356	4.90	0.0030
C2*ROW	8	0.60283912	0.07535489	6.66	0.0004
C3*ROW	8	0.32440799	0.04055100	3.59	0.0116
C4*ROW	8	0.13876129	0.01734516	1.53	0.2142
Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	6	0.25638292	0.04273049	3.78	0.0130
ROW	7	0.38831490	0.05547356	4.90	0.0030
C2*ROW	8	0.59754712	0.07469339	6.60	0.0004
C3*ROW	8	0.32649657	0.04081207	3.61	0.0113
C4*ROW	8	0.13876129	0.01734516	1.53	0.2142

Trend ANOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	22	1.79302842	0.08150129	19.38	0.0001
Error	33	0.13874820	0.00420449		
Corrected Total	55	1.93177662			

R-Square	C.V.	Root MSE	Y Mean
0.928176	6.404311	0.064842	1.012475

Dependent Variable: Y

Source	DF	Type I SS	Mean Square	F Value	Pr > F
TRT	6	0.27387545	0.04564591	10.86	0.0001
C1	1	0.09681321	0.09681321	23.03	0.0001
C2	1	0.53598746	0.53598746	127.48	0.0001
C3	1	0.22278336	0.22278336	52.99	0.0001
C5	1	0.13314475	0.13314475	31.67	0.0001
R1	1	0.27808763	0.27808763	66.14	0.0001
R2	1	0.02147675	0.02147675	5.11	0.0305
R3	1	0.04373966	0.04373966	10.40	0.0028
R5	1	0.02033078	0.02033078	4.84	0.0350
R6	1	0.01185195	0.01185195	2.82	0.1026
R7	1	0.01086024	0.01086024	2.58	0.1175
C1*R1	1	0.00973558	0.00973558	2.32	0.1376

C2*R3	1	0.01107563	0.01107563	2.63	0.1141
C3*R2	1	0.04705541	0.04705541	11.19	0.0021
R1*C4	1	0.04578624	0.04578624	10.89	0.0023
R2*C4	1	0.00916801	0.00916801	2.18	0.1492
C2*R1	1	0.02125631	0.02125631	5.06	0.0313

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	6	0.16044158	0.02674026	6.36	0.0002
C1	1	0.06777963	0.06777963	16.12	0.0003
C2	1	0.44309828	0.44309828	105.39	0.0001
C3	1	0.24999420	0.24999420	59.46	0.0001
C5	1	0.13222351	0.13222351	31.45	0.0001
R1	1	0.27808763	0.27808763	66.14	0.0001
R2	1	0.02147675	0.02147675	5.11	0.0305
R3	1	0.04373966	0.04373966	10.40	0.0028
R5	1	0.02033078	0.02033078	4.84	0.0350
R6	1	0.01185195	0.01185195	2.82	0.1026
R7	1	0.01086024	0.01086024	2.58	0.1175
C1*R1	1	0.00914040	0.00914040	2.17	0.1498
C2*R3	1	0.01580043	0.01580043	3.76	0.0611
C3*R2	1	0.04870965	0.04870965	11.59	0.0018
R1*C4	1	0.04431490	0.04431490	10.54	0.0027
R2*C4	1	0.01028565	0.01028565	2.45	0.1273
C2*R1	1	0.02125631	0.02125631	5.06	0.0313

Covariance Parameter Estimates (REML)

Cov Parm	Estimate
C1	0.00843481
C2	0.06534973
C3	0.03944736
C5	0.01928089
R1	0.03912510
R2	0.00246616
R3	0.00564660
R5	0.00230245
R6	0.00109118
R7	0.00094951
C1*R1	0.00559139
C2*R1	0.01769383
C2*R3	0.01540992
C3*R2	0.04762647
R1*C4	0.04172363
R2*C4	0.00559275
Residual	0.00421378

Least Squares Means

Effect	TRT	LSMEAN	Std Error	DF	t	Pr > |t|
TRT	1	1.03145832	0.02506657	33	41.15	0.0001
TRT	2	1.03632328	0.02409811	33	43.00	0.0001
TRT	3	1.08344910	0.02517848	33	43.03	0.0001
TRT	4	1.06286153	0.02574839	33	41.28	0.0001
TRT	5	0.95488139	0.02447435	33	39.02	0.0001
TRT	6	1.01891389	0.02524623	33	40.36	0.0001
TRT	7	0.89943749	0.02437852	33	36.89	0.0001

Row-column and interaction of regressions ANOVA

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	1.79923177	0.07196927	16.29	0.0001
Error	30	0.13254485	0.00441816		
Corrected Total	55	1.93177662			

R-Square	C.V.	Root MSE	Y Mean
0.931387	6.565027	0.066469	1.012475

Dependent Variable: Y

Source	DF	Type I SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	12.56	0.0001
COL	6	1.15907213	0.19317869	43.72	0.0001
TRT	6	0.11972355	0.01995392	4.52	0.0023
C1*R1	1	0.00957865	0.00957865	2.17	0.1513
R1*C2	1	0.01825578	0.01825578	4.13	0.0510
C2*R3	1	0.00785874	0.00785874	1.78	0.1923
C3*R2	1	0.04166095	0.04166095	9.43	0.0045
R1*C4	1	0.04499265	0.04499265	10.18	0.0033
R2*C4	1	0.00977442	0.00977442	2.21	0.1473

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ROW	7	0.38831490	0.05547356	12.56	0.0001
COL	6	1.01906239	0.16984373	38.44	0.0001
TRT	6	0.11791625	0.01965271	4.45	0.0025
C1*R1	1	0.00939749	0.00939749	2.13	0.1551
R1*C2	1	0.02030565	0.02030565	4.60	0.0403
C2*R3	1	0.01290053	0.01290053	2.92	0.0978
C3*R2	1	0.04269878	0.04269878	9.66	0.0041
R1*C4	1	0.04417127	0.04417127	10.00	0.0036
R2*C4	1	0.00977442	0.00977442	2.21	0.1473

Covariance Parameter Estimates (REML)

Cov Parm	Subject	Estimate
ROW		0.00729090
COL		0.02179930
C1*R1		0.00584283
R1*C2		0.01598859
C2*R3		0.01084891
C3*R2		0.04046662
R1*C4		0.04157133
R2*C4		0.00474734
SP(EXP)INTERCEPT		-0.00000000
Residual		0.00443729

Least Squares Means

Effect	TRT	LSMEAN	Std Error	DF	t	Pr > |t|
TRT	1	1.03279947	0.06849923	49	15.08	0.0001
TRT	2	1.04085965	0.06827608	49	15.24	0.0001
TRT	3	1.07188050	0.06898348	49	15.54	0.0001
TRT	4	1.05156492	0.06912807	49	15.21	0.0001
TRT	5	0.96546152	0.06879786	49	14.03	0.0001

TRT	6	1.02168355	0.06853511	49	14.91	0.0001
TRT	7	0.90307540	0.06835031	49	13.21	0.0001

Differences of Least Squares Means

Effect	TRT	_TRT	Difference	Std Error	DF	t	Pr > |t|
TRT	1	2	-0.00806018	0.03504670	49	-0.23	0.8191
TRT	1	3	-0.03908103	0.03618394	49	-1.08	0.2854
TRT	1	4	-0.01876545	0.03958021	49	-0.47	0.6375
TRT	1	5	0.06733794	0.03740107	49	1.80	0.0780
TRT	1	6	0.01111592	0.03811781	49	0.29	0.7718
TRT	1	7	0.12972407	0.03761943	49	3.45	0.0012
TRT	2	3	-0.03102085	0.03841115	49	-0.81	0.4232
TRT	2	4	-0.01070527	0.03929203	49	-0.27	0.7864
TRT	2	5	0.07539812	0.03608589	49	2.09	0.0419
TRT	2	6	0.01917610	0.03569990	49	0.54	0.5936
TRT	2	7	0.13778425	0.03651435	49	3.77	0.0004
TRT	3	4	0.02031558	0.03754102	49	0.54	0.5909
TRT	3	5	0.10641897	0.04097063	49	2.60	0.0124
TRT	3	6	0.05019695	0.03892509	49	1.29	0.2033
TRT	3	7	0.16880510	0.03807134	49	4.43	0.0001
TRT	4	5	0.08610340	0.03927030	49	2.19	0.0331
TRT	4	6	0.02988137	0.03847642	49	0.78	0.4411
TRT	4	7	0.14848952	0.03787633	49	3.92	0.0003
TRT	5	6	-0.05622202	0.03756983	49	-1.50	0.1409
TRT	5	7	0.06238613	0.03639929	49	1.71	0.0929
TRT	6	7	0.11860815	0.03565169	49	3.33	0.0017

Differences of Least Squares Means

Adjustment	Adj P
Tukey-Kramer	1.0000
Tukey-Kramer	0.9310
Tukey-Kramer	0.9991
Tukey-Kramer	0.5541
Tukey-Kramer	0.9999
Tukey-Kramer	0.0187
Tukey-Kramer	0.9831
Tukey-Kramer	1.0000
Tukey-Kramer	0.3749
Tukey-Kramer	0.9981
Tukey-Kramer	0.0074
Tukey-Kramer	0.9980
Tukey-Kramer	0.1492
Tukey-Kramer	0.8534
Tukey-Kramer	0.0010
Tukey-Kramer	0.3182
Tukey-Kramer	0.9862
Tukey-Kramer	0.0048
Tukey-Kramer	0.7455
Tukey-Kramer	0.6103
Tukey-Kramer	0.0260