Stat 8311, Fall 2006, general linear hypothesis

All the contrast functions in R/S-Plus *assume* that you want to define the space spanned by a factor to be orthogonal to the column of ones, J. This is evident from the examples below:

```
> set.seed(112)
> y <- rnorm(8)
> J2 <- c(1,1)
> # cell means parameterization
> (f <- factor( kronecker( 1:4, J2 )))
1  1  2  2  3  3  4  4
> (X <- data.frame(model.matrix(~-1+f)))
  f1 f2 f3 f4
1  1  0  0  0
2  1  0  0  0
3  0  1  0  0
4  0  1  0  0
5  0  0  1  0
6  0  0  1  0
7  0  0  0  1
8  0  0  0  1
> m0 <- lm(y ~ +1 + X$f4)  # didn't need to specify intercept
> m1 <- lm(y ~ -1 + X$f1+X$f2+X$f3+X$f4) # again, didn't need intercept
> anova(m0,m1)
Analysis of Variance Table

Model 1: y ~ +1 + X$f4
Model 2: y ~ -1 + X$f1 + X$f2 + X$f3 + X$f4
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1      6 10.8108
2      4  4.6866  2    6.1241 2.6135 0.1879
> anova(m1)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value Pr(>F)
X$f1       1 2.1823  2.1823  1.8625 0.2441
X$f2       1 3.0724  3.0724  2.6223 0.1807
X$f3       1 1.7017  1.7017  1.4524 0.2946
X$f4       1 1.5525  1.5525  1.3251 0.3138
Residuals  4 4.6866  1.1717
```

```
> # helmert parameterization
> (X <- data.frame(const=rep(1,8),kronecker(contr.helmert(4),J2)))
  const X1 X2 X3
1     1 -1 -1 -1
2     1 -1 -1 -1
3     1  1 -1 -1
4     1  1 -1 -1
5     1  0  2 -1
6     1  0  2 -1
7     1  0  0  3
8     1  0  0  3
> n0 <- lm(y ~ -1 + X$const + X$X3)
> n1 <- lm(y ~ -1 + X$const + X$X1 + X$X2 + X$X3)
> anova(n0,n1)
Analysis of Variance Table

Model 1: y ~ -1 + X$const + X$X3
Model 2: y ~ -1 + X$const + X$X1 + X$X2 + X$X3
  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1      6 10.8108
2      4  4.6866  2    6.1241 2.6135 0.1879
> anova(n1)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value Pr(>F)
X$const    1 0.0279  0.0279  0.0238 0.8848
X$X1       1 5.2167  5.2167  4.4524 0.1025
X$X2       1 0.9075  0.9075  0.7745 0.4285
X$X3       1 2.3568  2.3568  2.0116 0.2291
Residuals  4 4.6866  1.1717
> anova(update(n1,~-1 + X$const + X$X3 + X$X2 + X$X1))
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value Pr(>F)
X$const    1 0.0279  0.0279  0.0238 0.8848
X$X3       1 2.3568  2.3568  2.0116 0.2291
X$X2       1 0.9075  0.9075  0.7745 0.4285
X$X1       1 5.2167  5.2167  4.4524 0.1025
Residuals  4 4.6866  1.1717
```