Type I and Type II Anova

Type I (sequential) anova is given by the R command "anova(modl)". It shows how the RSS decreases as each predictor is added to the model. It changes if you order the predictors in the model differently.

Type II anova is given by the CAR command "Anova(modl)" It shows how the RSS would increase if each predictor in the model was removed, leaving the other predictors in. It does not change if you reorder the predictors in the model. In a regression, Type II gives the same tests you get from the t tests of the coefficients in the model. For example in the BGSgirls data:

```
lm(formula = Soma ~ WT2 + HT2 + WT9 + HT9 + WT18 + HT18)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.218252   1.741023   6.443 1.85e-08 ***
WT2          0.011173   0.056505   0.198  0.84389
HT2         -0.038392   0.026418  -1.453  0.15112
WT9          0.048933   0.019473   2.513  0.01455 *
HT9          0.002544   0.024249   0.105  0.91676
WT18         0.066650   0.009256   7.201 8.92e-10 ***
HT18        -0.054693   0.017007  -3.216  0.00205 **

> anova(fitter)      # Sequentail analysis of variance
Analysis of Variance Table

Response: Soma
          Df  Sum Sq Mean Sq F value    Pr(>F)
WT2        1  3.3041  3.3041 17.2649 9.981e-05 ***
HT2        1  1.4018  1.4018  7.3250  0.008740 **
WT9        1 17.6050 17.6050 91.9921 6.221e-14 ***
HT9        1  0.2488  0.2488  1.3002  0.258492
WT18       1  8.2224  8.2224 42.9647 1.186e-08 ***
HT18       1  1.9791  1.9791 10.3417  0.002054 **
Residuals 63 12.0566  0.1914

> Anova(fitter)
Anova Table (Type II tests)

Response: Soma
           Sum Sq Df F value    Pr(>F)
WT2        0.0075  1  0.0391  0.843893
HT2        0.4042  1  2.1119  0.151120
WT9        1.2084  1  6.3143  0.014548 *
HT9        0.0021  1  0.0110  0.916764
WT18       9.9232  1 51.8521 8.919e-10 ***
HT18       1.9791  1 10.3417  0.002054 **
Residuals 12.0566 63
```

Note that the Anova P values are identical to the P values of the regression summary. But only the last line of the anova matches that of the regression summary.

When you go to factors, there is no longer a simple connection between the "summary" table and either anova or Anova, but the same interpretation holds. For example a data set with two factors A having three levels and B having four gives:`anova(fitone)`

```
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
A          2 257.05 128.527 13.1876 4.469e-05 ***
B          3  25.69   8.563  0.8786   0.46074
A:B        6 150.91  25.151  2.5807   0.03391 *
Residuals 38 370.35   9.746

Anova(fitone)
Anova Table (Type II tests)
Response: y
          Sum Sq Df F value    Pr(>F)
A         240.14  2 12.3200 7.511e-05 ***
B          25.69  3  0.8786   0.46074
A:B       150.91  6  2.5807   0.03391 *
Residuals 370.35 38
---
lm(formula = y ~ B + A + B:A)

Residuals:
    Min      1Q  Median      3Q     Max
-7.0061 -1.4603  0.1868  1.4229  7.0433

> anova(fitthree)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq F value    Pr(>F)
B          3  42.60  14.200  1.4571   0.24158
A          2 240.14 120.071 12.3200 7.511e-05 ***
B:A        6 150.91  25.151  2.5807   0.03391 *
Residuals 38 370.35   9.746

> Anova(fitthree)
Anova Table (Type II tests)

Response: y
          Sum Sq Df F value    Pr(>F)
B          25.69  3  0.8786   0.46074
A         240.14  2 12.3200 7.511e-05 ***
B:A       150.91  6  2.5807   0.03391 *
Residuals 370.35 38
```

I reordered the factors A and B in the two fits. The interaction term is the same in all the anova and Anova, and the two Anovas are the same. But the Type I anovas are not; the contribution of each factor depends on whether it comes first or second the model fit.

This applies also to glms.