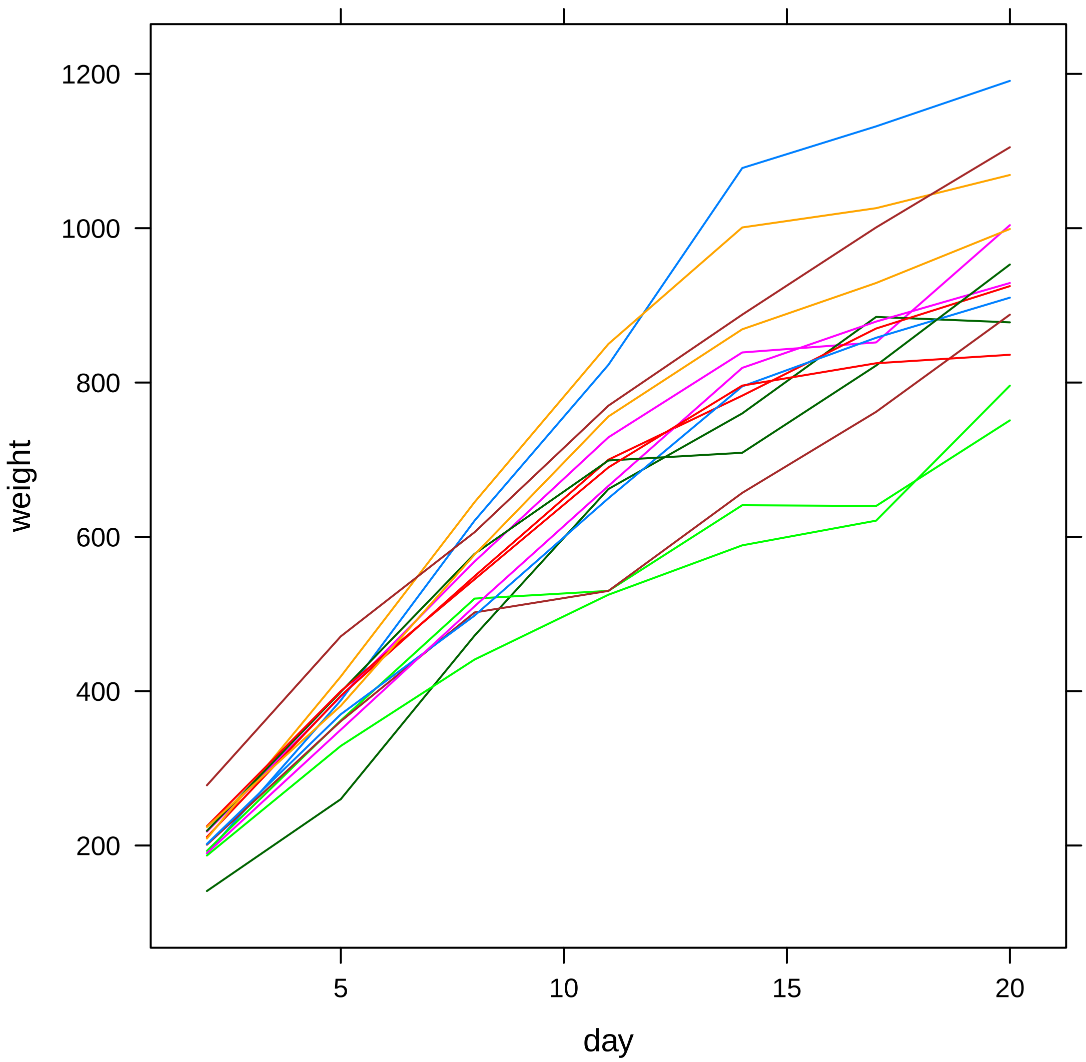
```
> sm <- read.delim("http://rem.ph.ucla.edu/rob/mld/data/tabdelimiteddata/smallmice.txt")
> head(sm, 10)

   Group id weight day
1      3 22    190   2
2      3 22    388   5
3      3 22    621   8
4      3 22    823  11
5      3 22   1078  14
6      3 22   1132  17
7      3 22   1191  20
8      3 23    218   2
9      3 23    393   5
10     3 23    568   8

> plot(xyplot(weight ~ day, group = id, type = "l", data = sm))
```



AR(1): Watch out...

Here are a bunch of ways to fit an `AR1` model.

- Using the form `1|id` assumes the observations are in order for each `id` and evenly spaced.

```
> ar1 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~1 |
+     id), data = sm)
> covonly(ar1)

Correlation Structure: AR(1)
 Formula: ~1 | id
 Parameter estimate(s):
     Phi
0.873738

Residual standard error: 93.79071
```

  If the data is mixed up, this won't do the right thing.

```
> smix <- sm[sample(1:nrow(sm)), ]
> head(smix)

   Group id weight day
20     3 24    885  17
67     3 31    699  11
89     4 34    589  14
28     3 25    925  20
10     3 23    568   8
66     3 31    578   8

> ar1b <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~1 |
+     id), data = smix)
> covonly(ar1b)

Correlation Structure: AR(1)
 Formula: ~1 | id
 Parameter estimate(s):
      Phi
0.6290928

Residual standard error: 98.5571
```

- Using the form `day|id` uses `day` to set the order for each `id`, BUT, it must be integer-valued and each integer up to the max must be present for at least one `id`. So this doesn't work here.

```
> ar2 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~day |
+     id), data = sm)
> covonly(ar2)

Correlation Structure: ARMA(1,0)
 Formula: ~day | id
 Parameter estimate(s):
Phi1
   0

Residual standard error: 97.0382
```

We're missing some days; what to do?

- Using `corAR1` allows `day` to be continuous. Order doesn't matter

```
> ar3 <- gls(weight ~ poly(day, 2), correlation = corCAR1(form = ~day |
+     id), data = sm)
> covonly(ar3)

Correlation Structure: Continuous AR(1)
 Formula: ~day | id
 Parameter estimate(s):
      Phi
0.9560055

Residual standard error: 93.7907

> ar3b <- gls(weight ~ poly(day, 2), correlation = corCAR1(form = ~day |
+     id), data = smix)
> covonly(ar3b)

Correlation Structure: Continuous AR(1)
 Formula: ~day | id
 Parameter estimate(s):
      Phi
0.9560055

Residual standard error: 93.7907
```

- Or, make an integer valued variable to set the order. (This works on the mixed up data too; not shown.)

```
> sm$dayF <- factor(sm$day)
> head(sm$dayF, 9)

[1] 2  5  8  11 14 17 20 2  5
Levels: 2 5 8 11 14 17 20

> sm$dayF <- as.numeric(sm$dayF)
> head(sm$dayF, 9)

[1] 1 2 3 4 5 6 7 1 2

> ar4 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~dayF |
+     id), data = sm)
> covonly(ar4)

Correlation Structure: AR(1)
 Formula: ~dayF | id
 Parameter estimate(s):
     Phi
0.873738

Residual standard error: 93.79071
```

- Here, we could also do this by arithmetic on the original day value.

```
> sm$day3 <- (sm$day + 1)/3
> ar5 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~day3 |
+     id), data = sm)
> covonly(ar5)

Correlation Structure: AR(1)
 Formula: ~day3 | id
 Parameter estimate(s):
     Phi
0.873738

Residual standard error: 93.79071
```

This works on the mixed up data too.

```
> smix$day3 <- (smix$day + 1)/3
> ar5 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~day3 |
+     id), data = smix)
> covonly(ar5)

Correlation Structure: ARMA(1,0)
 Formula: ~day3 | id
 Parameter estimate(s):
    Phi1
0.873738

Residual standard error: 93.79071
```

Why doesn't the `CAR1` method agree? It's using the actual day as the unit; the others are using three days as the unit.

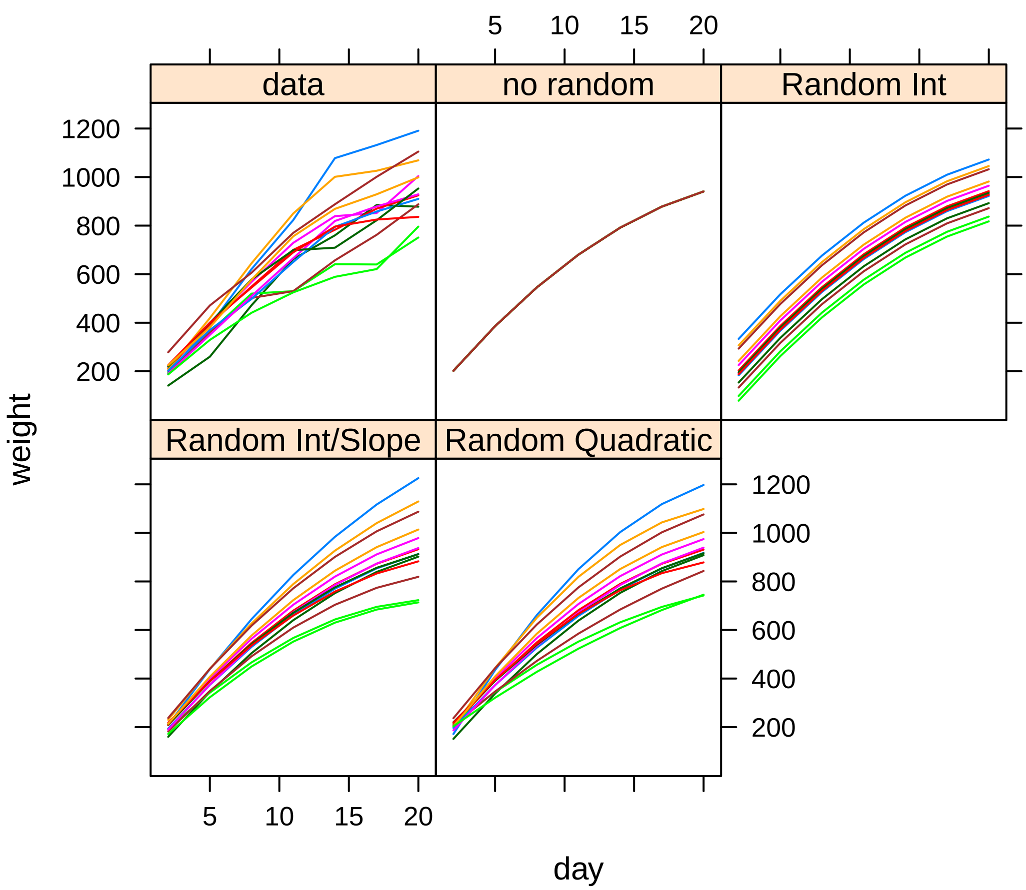
```
> 0.9560055^3

[1] 0.8737379
```

Fitting a bunch of models

```
> m0 <- gls(weight ~ poly(day, 2), data = sm)
> m1 <- lme(weight ~ poly(day, 2), random = ~1 | id, data = sm)
> m2 <- lme(weight ~ poly(day, 2), random = ~day | id, data = sm)
> m3 <- lme(weight ~ poly(day, 2), random = ~poly(day, 2) | id,
+     data = sm)
> m.ar1 <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~day3 |
+     id), data = sm)
> m1.ar1 <- lme(weight ~ poly(day, 2), random = ~1 | id, correlation = corAR1(form = ~day3 |
+     id), data = sm)
> anova(m0, m1, m2, m3, m.ar1, m1.ar1, test = FALSE)
```

```
       Model df      AIC      BIC    logLik
m0         1  4 1151.453 1161.669 -571.7266
m1         2  5 1095.651 1108.421 -542.8257
m2         3  7 1032.969 1050.846 -509.4844
m3         4 10 1027.690 1053.229 -503.8452
m.ar1      5  5 1025.619 1038.388 -507.8095
m1.ar1     6  6 1027.619 1042.942 -507.8095
```



Some more models...

```
> m.ar1.w <- gls(weight ~ poly(day, 2), correlation = corAR1(form = ~day3 |
+     id), weights = varIdent(form = ~1 | day), data = sm)
> covonly(m.ar1.w)

Correlation Structure: AR(1)
 Formula: ~day3 | id
 Parameter estimate(s):
      Phi
0.8963247
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | day
 Parameter estimates:
       2        5        8       11       14       17       20
1.000000 1.718826 2.097826 3.244799 3.903767 3.923631 3.487493

Residual standard error: 31.80178

> anova(m3, m.ar1, m.ar1.w)

        Model df       AIC      BIC    logLik   Test  L.Ratio p-value
m3          1 10 1027.6905 1053.229 -503.8452
m.ar1       2  5 1025.6189 1038.388 -507.8095 1 vs 2  7.92847  0.1602
m.ar1.w     3 11  993.2314 1021.324 -485.6157 2 vs 3 44.38752  <.0001

> m2.ar1 <- lme(weight ~ poly(day, 2), random = ~1 | id, correlation = corAR1(form = ~day3 |
+     id), data = sm)
> covonly(m2.ar1)

Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:  0.02137176 93.78968

Correlation Structure: AR(1)
 Formula: ~day3 | id
 Parameter estimate(s):
      Phi
0.8737346

> m2.w <- lme(weight ~ poly(day, 2), random = ~1 | id, weights = varIdent(form = ~1 |
+     day), data = sm)
> covonly(m2.w)

Random effects:
 Formula: ~1 | id
        (Intercept) Residual
StdDev:    57.87724 47.56953

Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | day
 Parameter estimates:
        2         5         8        11        14        17        20
1.0000000 0.7687003 0.1095683 1.2176693 1.9142864 2.0745132 1.6855478
```

```
> anova(m3, m.ar1.w, m2.ar1, m2.w)

        Model df       AIC      BIC    logLik   Test  L.Ratio p-value
m3          1 10 1027.6905 1053.229 -503.8452
m.ar1.w     2 11  993.2314 1021.324 -485.6157 1 vs 2 36.45905  <.0001
m2.ar1      3  6 1027.6189 1042.942 -507.8095 2 vs 3 44.38752  <.0001
m2.w        4 11 1084.8594 1112.952 -531.4297 3 vs 4 47.24047  <.0001

> m3.ar1 <- lme(weight ~ poly(day, 2), random=~poly(day, 2)|id, correlation=corAR1(form=~day3|id), data=
Error in lme.formula(weight ~ poly(day, 2), random = ~poly(day, 2) | id,  :
  nlminb problem, convergence error code = 1
  message = iteration limit reached without convergence (9)
```

Some unstructured models

```
> m.uncor <- gls(weight ~ poly(day, 2), correlation = corSymm(form = ~1 |
+     id), data = sm)
> covonly(m.uncor)

Correlation Structure: General
 Formula: ~1 | id
 Parameter estimate(s):
 Correlation:
  1      2      3      4      5      6
2  0.978
3  0.897  0.932
4  0.647  0.701  0.854
5 -0.101  0.015  0.237  0.607
6  0.185  0.262  0.424  0.730  0.857
7  0.460  0.538  0.667  0.806  0.685  0.841

Residual standard error: 100.6841

> m.un <- gls(weight ~ poly(day, 2), correlation = corSymm(form = ~1 |
+     id), weights = varIdent(form = ~1 | day), data = sm)
> covonly(m.un)

Correlation Structure: General
 Formula: ~1 | id
 Parameter estimate(s):
 Correlation:
  1     2     3     4     5     6
2 0.919
3 0.496 0.732
4 0.303 0.506 0.874
5 0.178 0.420 0.814 0.933
6 0.233 0.418 0.750 0.901 0.937
7 0.349 0.531 0.811 0.866 0.887 0.919
Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | day
 Parameter estimates:
       2        5        8       11       14       17       20
1.000000 1.572825 1.971261 3.414574 4.448796 4.456199 3.992665

Residual standard error: 30.15103
```

```
> anova(m.ar1.w, m.uncor, m.un)

         Model df       AIC      BIC    logLik   Test  L.Ratio p-value
m.ar1.w      1 11  993.2314 1021.324 -485.6157
m.uncor      2 25 1033.3396 1097.187 -491.6698 1 vs 2 12.10818  0.5976
m.un         3 31 1007.0444 1086.215 -472.5222 2 vs 3 38.29521  <.0001
```