# Mixed Models

Pedro Victor Brasil Ribeiro

2021-11-22 - Last changed in 2021-12-05

## Longitudinal Data

Longitudinal data, also called as panel data, is data that is collected through a series of repeated observations of the same subject overtime. Longitudinal experiments planning concern the observation of one or more variables in the same subject on different ocassions or condicion of evaluation, time is a commom factor used in most of experiment, but it can be use on distances among other factors, but in the present work it will be implied that the subjects measures is taken over time.

Given that longitudinal data are measures of the same subject taken in a systematic way, is expected not null correlations between the measures, especially the one taken in consecutive. Furthermore is expected heterocedasticidy.

About the data structure is expected 3 characteristics, but have in mind that in a real experiement ambient, not necessarily those characteristics can actually be hold:

- Regular (in respect to time): The interval of time between one measure and other are the same;
- Balanced (in respect to time): All observations are taken in the same time, on the same condictions in all subjects;
- Complete: No lost observations.

Table 1: Basic structure of balanced and complete longitudinal data

| Grupo ou Tratamento | Unidade Experimental | Condições de Avaliação | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | $\cdots$ | t |
| 1 | 1 | $y_{111}$ | $y_{112}$ | $\cdots$ | $y_{11t}$ |
| | 2 | $y_{121}$ | $y_{122}$ | $\cdots$ | $y_{12t}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | $n_1$ | $y_{1n_11}$ | $y_{1n_12}$ | $\cdots$ | $y_{1n_1t}$ |
| 2 | 1 | $y_{211}$ | $y_{212}$ | $\cdots$ | $y_{21t}$ |
| | 2 | $y_{221}$ | $y_{222}$ | $\cdots$ | $y_{22t}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | $n_2$ | $y_{2n_21}$ | $y_{2n_22}$ | $\cdots$ | $y_{2n_2t}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| g | 1 | $y_{g11}$ | $y_{g12}$ | $\cdots$ | $y_{g1t}$ |
| | 2 | $y_{g21}$ | $y_{g22}$ | $\cdots$ | $y_{g2t}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| | $n_g$ | $y_{gn_g1}$ | $y_{gn_g2}$ | $\cdots$ | $y_{gn_gt}$ |

# Mixed Model

For the analysis of longitudinal data to take into account the correlation between measurements taken in the same experimental unit, the adjustment of regression models with the inclusion of random effects can be considered, and it's oftenly used in these kind of cases. The usual models are:

- Linear Mixed Models;
- Generalized Linear Mixed Models;
- Non-Linear Mixed Models.

In this we will speak specifically about Linear Mixed Models (LMM) since is the most common model used in this kind of situation. So we'll use LMM to fit the mean profile of the subjects.

As usually done in Linear Models, the LMM is use to make the inference on the populacional mean of the populacion, but in the LMM is used the y conditioned in u, in other words E[Y|u]. For a LMM $X\beta$ represent the fixed effects, but is added a random effect represented by Zu, where just as X is the knew model matriz, and u the random variables used to define the condicional model, expressed as:

$$E[Y|u] = X\beta + Zu + \varepsilon \tag{1}$$

Where from the model 1 is usually assumed that $b_i \sim N_q(0, G)$ and $\varepsilon \sim N_{m_i}(0, R_i)$, where $m_i$ is the number of instants that the subject is measured.

Table 2: Summary of the dimensions of each component

| Componente | Dimensão |
|---|---|
| y | $N \times 1$ |
| X | $N \times p$ |
| $\beta$ | $p \times 1$ |
| Z | $N \times nq$ |
| b | $nq \times 1$ |
| $\varepsilon$ | $N \times 1$ |
| $\Gamma$ | $nq \times nq$ |
| R | $N \times N$ |
| $\Omega$ | $N \times N$ |
| $\theta$ | $t \times t$ |

WHere in a more techinicall manner the modelo 1, which $y = [y_1^T, \cdots, y_n^T](N \times 1)$, with $N = \sum_{i=1}^{n} n_i$, have the fixed part and the random part, the parameter of the model can be write in a matter where $b \sim N_{nq}(0, \Gamma(\theta))$, where $\Gamma(\theta) = I_n \otimes G(\theta)$ is independent of $\varepsilon \sim N_N(0, R(\theta)), R(\theta) = \otimes_{i=1}^{n} R_i(\theta)$, with give the final distribuition of y as $y \sim N_N(X\beta, \Omega(\theta))$, where $\Omega(\theta) = Z\Gamma(\theta)Z^T + R(\theta)$.

Which show us that the fixed effect affect only the mean of y, when the random affect affect the variance of y.

# Structure of the covariance

The greater part of the effort of fit a Mixed Linear Model is to choose the structure of the covariance, which greatly affect the goodness-of-fit of the fitted model.

Theorically speaking those structure can be used on $\Omega(\theta), R(\theta)$ and $G(\theta)$, but is usually used in the variance of y ($\Omega(\theta)$).

Some exemples of covariante structure, considering $n_i = 4$:

## Uniform structure

$[\theta = (\sigma^2, \tau)^T]$

$$\begin{bmatrix} \sigma^2+\tau & \tau & \tau & \tau \\ \tau & \sigma^2+\tau & \tau & \tau \\ \tau & \tau & \sigma^2+\tau & \tau \\ \tau & \tau & \tau & \sigma^2+\tau \end{bmatrix}$$

## AR(1) structure

$[\theta = (\sigma^2, \phi)^T]$

$$\begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

## ARMA(1,1) structure

$[\theta = (\sigma^2, \phi, \gamma)^T]$

$$\begin{bmatrix} 1 & \gamma & \phi\gamma & \gamma\phi^2 \\ \gamma & 1 & \gamma & \gamma\phi \\ \gamma\phi & \gamma & 1 & \gamma \\ \gamma\phi^2 & \gamma\phi & \gamma & 1 \end{bmatrix}$$

## Antidependence structure of 1st order

$[\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \rho_1, \rho_2, \rho_3, \rho_4)^T]$

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_1\sigma_2\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_1\sigma_3\rho_1\rho_2 & \sigma_2\sigma_3\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_3 \\ \sigma_1\sigma_4\rho_1\rho_2\rho_3 & \sigma_2\sigma_4\rho_2\rho_3 & \sigma_3\sigma_4\rho_3 & \sigma_4^2 \end{bmatrix}$$

## Toeplitz structure

$[\theta = (\sigma^2, \sigma_1, \sigma_2, \sigma_3)^T]$

$$\begin{bmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{bmatrix}$$

### Hetorogenally uniform structure

$[\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \rho)^T]$

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$$

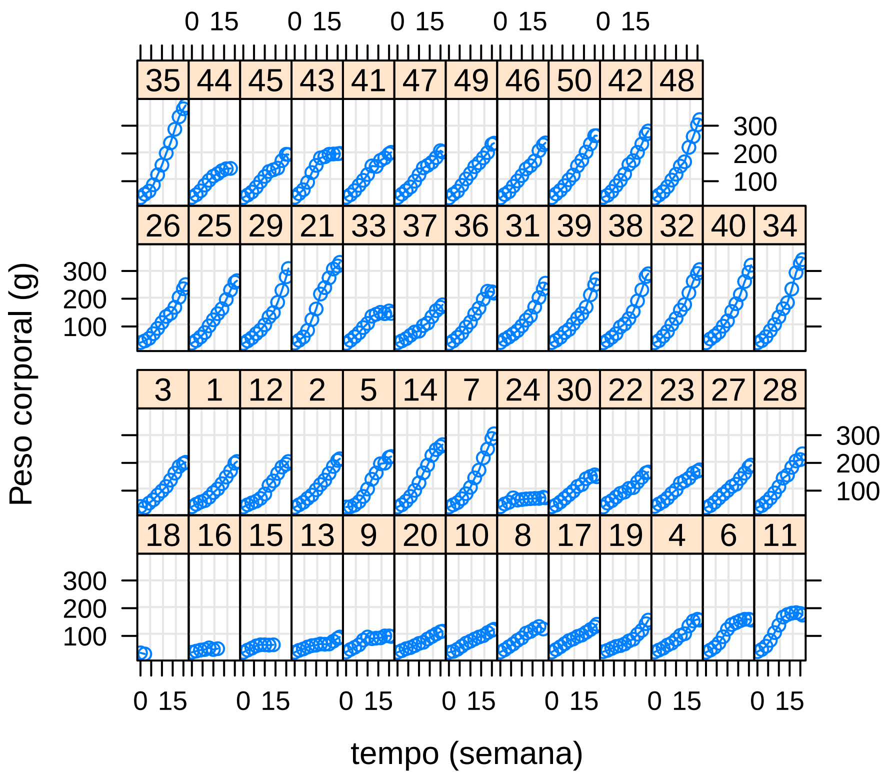
### Not structured

$[\theta = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_{12}, \sigma_{13}, \sigma_{14}, \sigma_{23}, \sigma_{24}, \sigma_{34})^T]$

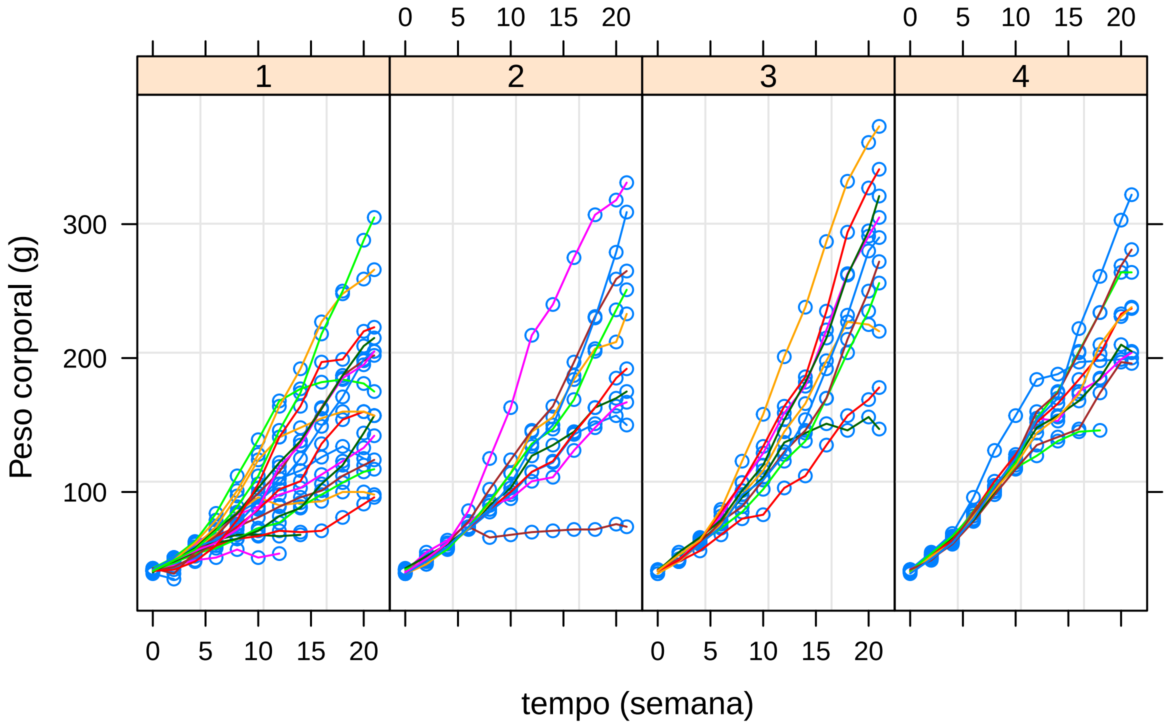
$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

## Dados Dieta - Frango

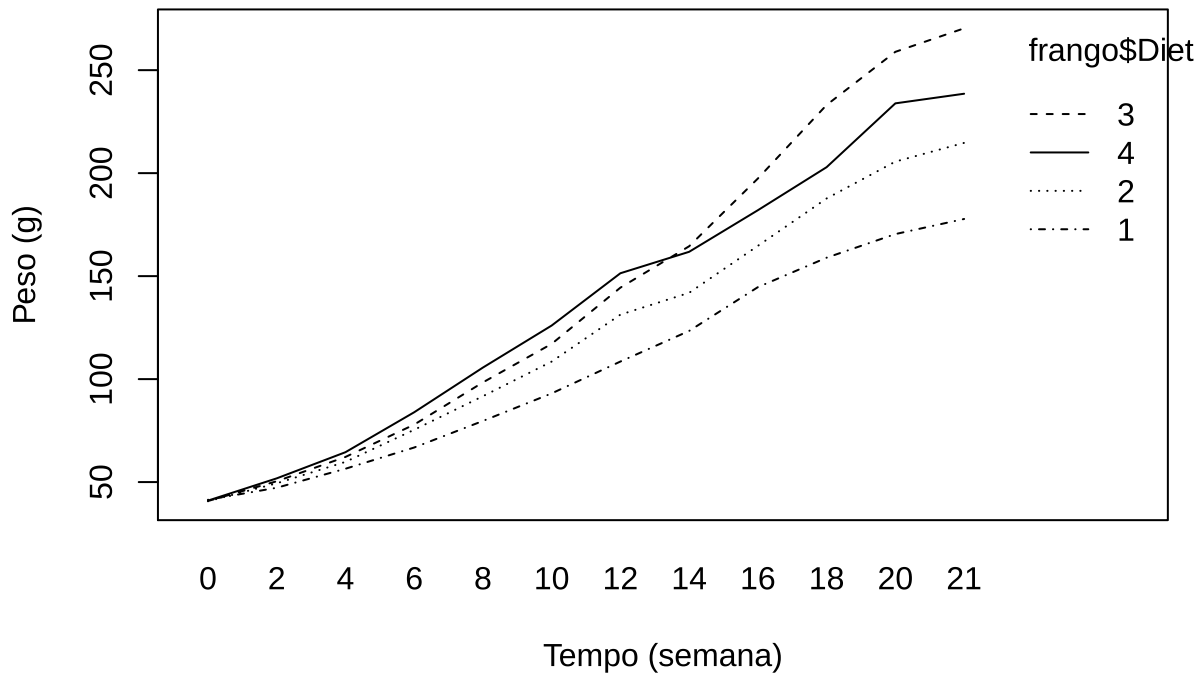
```
data("ChickWeight")

frango <- groupedData(
  weight~Time|Chick, outer = ~Diet, data = ChickWeight,
  order.groups = F,
  labels=list(x="tempo (semana)", y="Peso corporal (g)")
)

plot(frango,between = list(y = c(0, 0.5, 0)))
```




```
plot(frango, outer = T, key=FALSE) # key omite a legenda
```




```
interaction.plot(frango$Time,frango$Diet,frango$weight,ylab="Peso (g)", xlab="Tempo (semana)")
```




Por tanto podemos observar que os perfis individuais apresentados para cada frango diferem entre si, apresentando um perfil médio o qual apresenta uma clara modificação na variância, quando observado o peso em relação ao tempo, perfil o qual, varia de acordo com cada dieta utilizada.

```
mod1 <- lme(
  fixed = weight~Time+Diet+Time:Diet,
  random = ~1,
  data = frango,
  control = lmeControl(opt="optim")
)
mod2 <- lme(
  fixed = weight~Time+Diet+Time:Diet,
  random = ~ Diet,
  data = frango,
  control = lmeControl(opt="optim")
)
mod3 <- lme(
  fixed = weight~Time+Diet+Time:Diet,
  random = ~Time,
  data = frango,
  control = lmeControl(opt="optim")
)
mod4 <- lme(
  fixed = weight~Time+Diet+Time:Diet,
  random = ~Time + Diet,
  data = frango,
```

```
  control = lmeControl(opt="optim")
)
mod5 <- lme(
  fixed = weight~Time+Diet+Time:Diet,
  random = ~Time + Diet + Time:Diet,
  data = frango,
  control = lmeControl(opt="optim")
)
```

```
coef(mod3) %>%
  as_tibble() %>%
  kable(format = "latex", booktabs = TRUE)
```

| (Intercept) | Time | Diet2 | Diet3 | Diet4 | Time:Diet2 | Time:Diet3 | Time:Diet4 |
|---|---|---|---|---|---|---|---|
| 33.94980 | 6.0821404 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 48.91139 | 0.9275944 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 46.57686 | 2.0830318 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 46.01675 | 2.1230021 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 45.12966 | 3.1065073 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 41.35582 | 3.5399976 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 40.41212 | 3.9860585 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 38.24910 | 5.1860977 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 39.21681 | 4.7721350 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 37.07350 | 4.7542116 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 34.17914 | 6.0132306 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 33.79654 | 6.9839217 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 30.60247 | 8.5171166 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 26.67009 | 8.2613885 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 28.20287 | 7.7521221 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 26.75626 | 8.1371705 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 26.04689 | 8.6191198 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 21.76361 | 9.7336397 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 16.21147 | 12.1820996 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 11.97323 | 12.7796153 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 56.59527 | -0.9669800 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 42.12084 | 3.7099195 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 42.22226 | 3.7409063 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 39.76045 | 4.5887702 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 37.34690 | 4.9138575 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 30.21870 | 7.2899561 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 28.89805 | 7.5706638 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 25.29045 | 8.9128557 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 21.31593 | 9.4732140 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 12.77832 | 13.5369374 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 51.29844 | 1.3458263 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 48.18251 | 1.3897275 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 38.21537 | 5.0369531 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 37.82660 | 4.7000759 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 35.63375 | 5.4100898 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 31.44800 | 6.5930090 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 28.25827 | 8.0676981 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 27.36432 | 8.2114639 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 22.51005 | 9.7015956 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 15.80988 | 12.3136613 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 43.34374 | 3.3758280 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 39.27597 | 4.3397792 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 38.09903 | 6.0057850 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 38.02867 | 5.0974285 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 37.27198 | 5.1875901 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 33.09317 | 6.4797517 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 32.86842 | 6.2846650 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 27.98019 | 7.9168952 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 26.32363 | 8.2831550 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |
| 20.26239 | 9.7992229 | -5.021124 | -15.40439 | -1.742559 | 2.332126 | 5.145861 | 3.254743 |