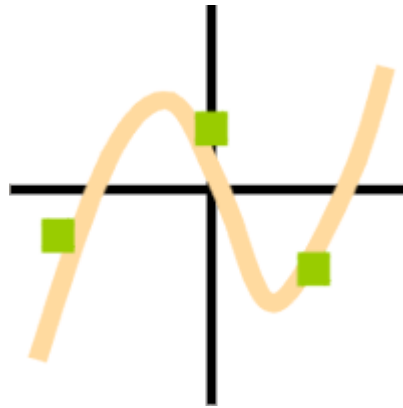


Manual for Regression 2003

(Version 1.0)



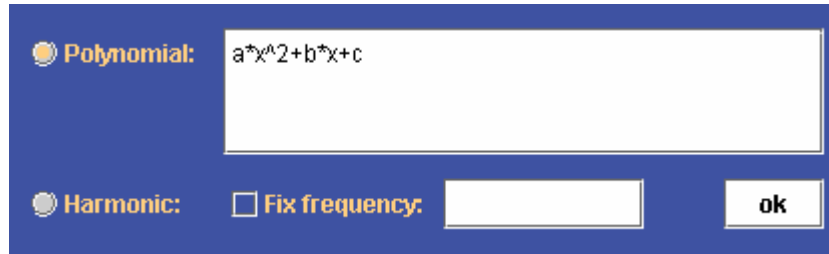
Abstract

Regression 2003 is a program for regression analysis including n-dimensional polynomial regression and one-dimensional harmonic regression. It was designed to be as flexible and easy to use as possible to fulfill its purpose as a tool for teaching. Regression 2003 was developed by Andreas Keilhauer (Keilhauer@gmx.de) as a programming practical for Prof. Leha's department for stochastic processes at the University of Passau.

Polynomial Regression

1. Problem input

Select "Polynomial" in the input area and enter an arbitrary polynomial:



Polynomial:

Harmonic: Fix frequency:

ok

A polynomial has to be given in expanded form (see above). It is absolutely necessary that every term (e.g.: x^3 , x_1*x_2) has exactly one variable coefficient that can be optimized. That means that a polynomial like " $x^2+b*x+c$ " is not allowed. Some examples of allowed polynomials are: " $x*b+d*x^2+a$ ", " $m*x+b$ ", " $a*x_1*x_2+c$ " and " $x*a_1+a$ ".

2. Mark rows

We are doing n-dimensional polynomial regression here, so the first n+1 columns of the **data table** have to be considered. In any case use x_1, \dots, x_n as variables or simply x for one- dimensional polynomials (see also: **mark rows**).

In our example only the first two columns are of any importance.

X	Y	AutoID	swing
0	1	1	1
1	4	2	0
2	4	3	-1
3	6	4	0
4	15	5	1
5	1	6	0

In general the first column contains the indices and the following n columns contain the values of x_1, \dots, x_n . To change the considered columns you can **move columns**.

3. Press the OK-Button in the input area.

Now the following solution and its squareloss is displayed in the solution area:



Solution:

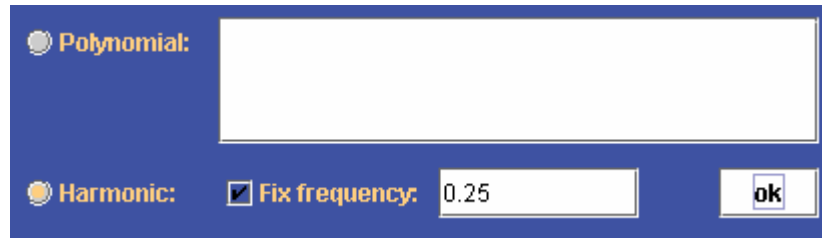
Squareloss:

By then, the plotter will already have drawn the solution and will have scaled itself to the data automatically.

Harmonic Regression

1. Problem input

Select "Harmonic" in the input area and decide whether you want to fix the frequency of the harmonic or not.



Polynomial:

Harmonic: Fix frequency:

In our example we fixed the frequency to 0.25 . You are also allowed to enter any expression like "1/4" or "(1/2)^2" (see parser documentation for details).

2. Mark rows

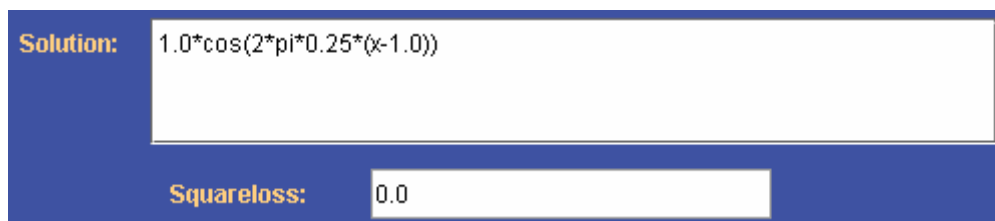
We are doing one dimensional harmonic regression here, so only the first two columns are of any importance (see also: **mark rows**).

AutoID	swing	X	Y
1	1	0	1
2	0	1	4
3	-1	2	4
4	0	3	6
5	1	4	15
6	0	5	1

The first column of the **data table** represents the index set (must be a continuous intervall) and the second one contains the x-values. To change the considered columns you can **move columns**.

3. Press the OK-Button in the input area.

Now the following solution and its squareloss is displayed in the solution area:



Solution:

Squareloss:

By then, the plotter will already have drawn the solution and will have scaled itself to the data automatically.

Parser documentation

Throughout the whole program, the JEP (Java Evaluation Parser) is used to parse arithmetic expressions. See <http://www.singularsys.com/jep/> for details.

Operators

All common arithmetic operators are supported. Boolean operators are also fully supported. Boolean expressions are evaluated to be either 1 or 0 (true or false respectively).

Power	\wedge
Boolean Not	!
Unary Plus, Unary Minus	+x, -x
Modulus	%
Division	/
Multiplication	*
Addition, Subtraction	+, -
Less or Equal, More or Equal	<=, >=
Less Than, Greater Than	<, >
Not Equal, Equal	!=, ==
Boolean And	&&
Boolean Or	

Functions

Sine	sin()
Cosine	cos()
Tangent	tan()
Arc Sine	asin()
Arc Cosine	acos()
Arc Tangent	atan()
Hyperbolic Sine	sinh()
Hyperbolic Cosine	cosh()
Inverse Hyperbolic Sine	asinh()
Inverse Hyperbolic Cosine	acosh()
Inverse Hyperbolic Tangent	atanh()
Natural Logarithm	ln()
Logarithm base 10	log()
Angle	angle()
Absolute Value / Magnitude	abs()
Random number (between 0 and 1)	rand()
Modulus	mod()
Square Root	sqrt()

Appendix

Add a Column



After pressing this button, a dialog asking for a name for the new column will appear. Enter any name and the new column will be attached to the **DataTable**.

Add Rows



After pressing this button, a dialog asking for the number of rows will appear. Enter a natural number n and n rows will be attached to the DataTable.

Data Table

This component is responsible for displaying the data. You can **open data** from a file or **enter data** yourself. It is also possible to **save data** to an arbitrary file. Moreover, you can transmit the first two columns of the data table to the **Graph Drawer** by using the **Plot Data Points** button.

Draw the result function



This draws the function (model) described in the textbox close-by and calculates its squareloss w.r.t. the data.

Enter Data

Before you can enter data, you probably need to **insert rows** or also **insert columns**. After doing so, there will be empty cells for your data. To enter an arbitrary floating point number into a cell of the **data table**, simply click on a cell, delete the value currently in the marked cell (e.g. by using the Backspace Button) and then type in any floating point number like "2.71828", "1.7E10" or "12345". After that, press the Return Button. The new value is now entered into the data table. In the following example we are about to enter "23.7" into the marked cell.

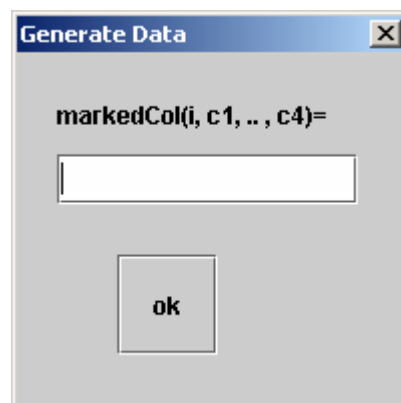
AutoID	X	Y	swing
1	0	1	1
2	1	4	0
3	2	4	-1
4	3	6	0
5	4	15	1
6	5	23.7	0

If you want to enter something more complex like "sin(e)", "1/7" or "4*pi", click on the cell and use **Generate Data** to achieve this.

Generate Data



Mark the rows and the column of the cells the data is to be generated for. Now press the button above and a dialog like the following will appear:



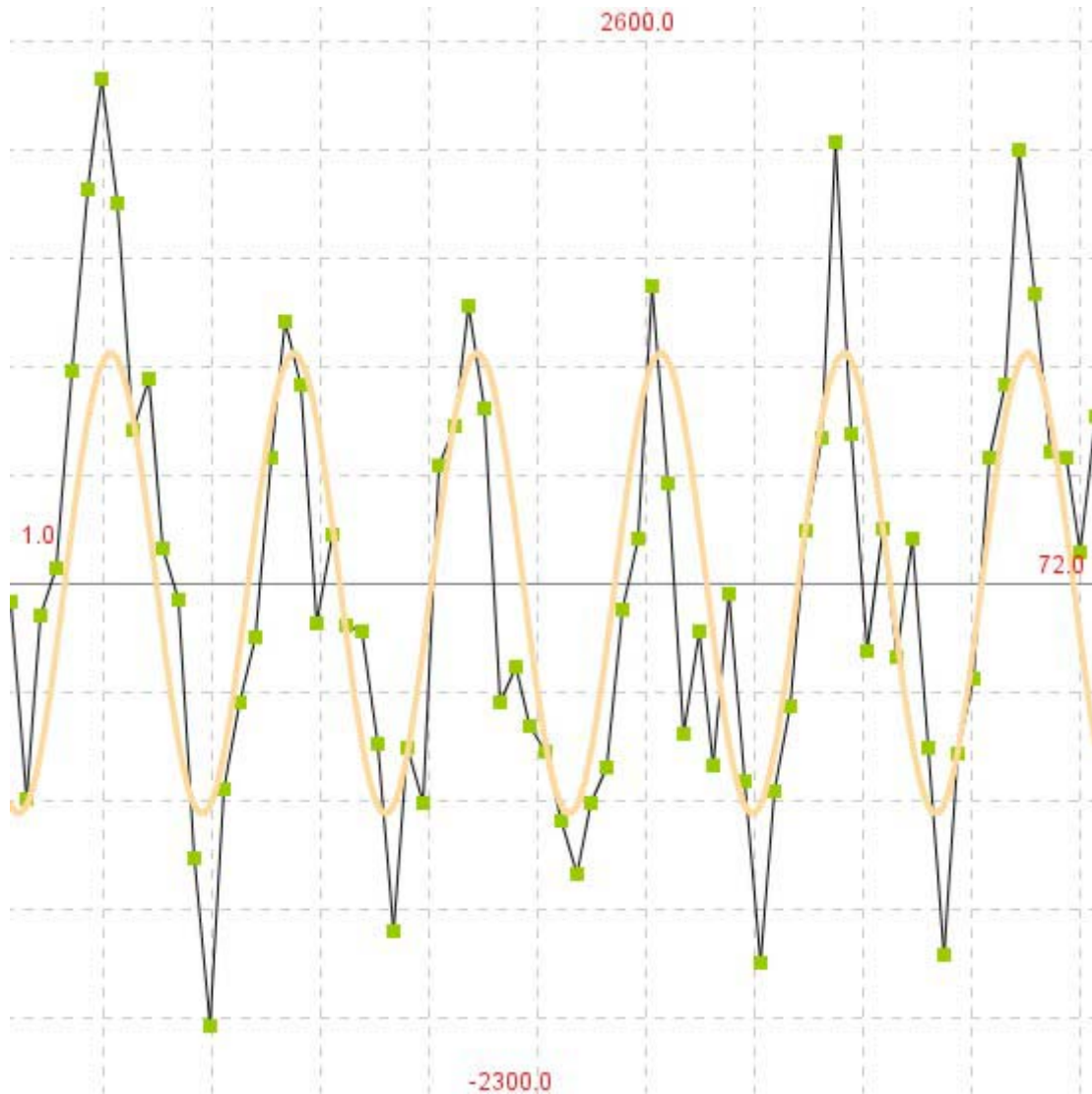
You are asked to enter a function that generates the data for all **marked rows** $R \subseteq \{1, \dots, \text{numberOfRows}\}$ in a **marked column** c . For all $i \in R$ the value in the i -th row of column c is set to the value of $\text{markedCol}(i, c_1, \dots, c_n)$, where c_1, \dots, c_n are the values in the i -th row of column $1, \dots, n$.

An example input could be: $i+2*c_1-\cos(c_4)$ (see parser docs for details)

Graph Drawer

This component **draws the result function, plots data points** and **lines between the datapoints**. It is also possible to determine the size of the grids in x and y-direction in the **Graph Drawer Settings** or to display the coordinates of the Graph Drawer by pressing the Shift Button and moving over it. Moreover, you can save the content of the Graph Drawer as an image file using the **Graph Drawer Image Export**.

The following example shows the Graph Drawer displaying a set of data points with lines between them and also a harmonic that has been laid through the data points with least square loss.



Graph Drawer Settings

To determine how the grid of the **Graph Drawer** should be displayed choose Settings → Graph Drawer. You can **save settings** to make them standard for the next time you use the program.

Graph Drawer Image Export

Choose File → Save Image to save the current **Graph Drawer** display as a png file.

Lines between Data Points



Pressing this button will cause the **Graph Drawer** to draw lines from one data point to the next one. Therefore their order in the data table, at the time of the last transmission

to the **Graph Drawer**, will be used. After pressing the button displayed above, its icon will change to the one below.



Pressing this button will cause the Graph Drawer not to draw lines between the data points. After clicking on it, its icon will change back.

This setting can be saved using the **save settings** option.

Mark Rows

A single row can be marked by simply clicking on one cell of this row. All rows can be marked by pressing Ctrl and A simultaneously.

More than one row can be marked like the following:

- Hold the Ctrl Button and click on all rows you want to choose one after another.
- To mark an area click on the first and the last row of the area while keeping the Shift Button pressed.
- Combine the ways above by keeping the Ctrl Button pressed to mark a new single selection or the beginning of a new area selection.

In that way, selections like the following are possible:

AutoID	X	Y	swing
1	0	1	1
2	1	4	0
3	2	4	-1
4	3	6	0
5	4	15	1
6	5	1	0

Mark a Column

A column c is marked if the editing focus is on a cell of c. Mark a column by clicking on any cell of the column. In the following example, column Y is marked.

AutoID	X	Y	swing
1	0	1	1
2	1	4	0
3	2	4	-1
4	3	6	0
5	4	15	1
6	5	1	0

Model Deviation Data



After pressing this button, a dialog will ask you to label the column that will contain the model deviation data. After that, the deviation of the data in the i -th row from the current model is inserted into the new column as the i -th entry. This column is now attached to the DataTable.

Move Columns

In order to change the considered columns for a regression analysis, you can move them to the desired position via drag and drop. In the following example column Y will be moved to the second position after you have released the mouse button.

AutoID		Y	X	swing
1		1	0	1
2		4	1	0
3		4	2	-1
4		6	3	0
5		15	4	1
6		1	5	0

Open Data

To open data from a file use File → Open. Alternatively, you can open one of the recently used files. Simply choose File → "Open Recent" and click on the file you want to open.

Plot Data Points



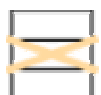
Pressing this button will cause the **Graph Drawer** to consider the most left column as x-coordinates and the second most left column as y-coordinates of points in \mathbb{R}^2 . Afterwards, it will plot the points that are related to the **marked rows**.

Remove Columns



Press this button to cause the **Data Table** to remove the **marked column**.

Remove Rows



Press this button to cause the **Data Table** to remove the **marked rows**.

Rounding

To determine how the system should round numbers choose Settings → Rounding. You can **save settings** to make them standard for the next time you use the program.

Save Data

To save data to a file use File → Save/"Save as...", where "Save" will write the data to the file currently used and "Save as..." will offer you to choose a file.

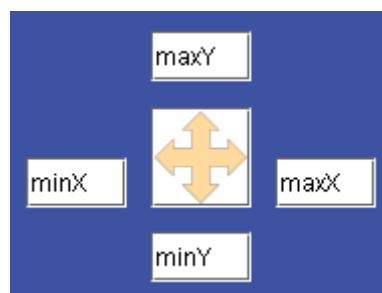
Save Settings

If you want to save the current settings as the standard for the next program, start choose Settings → "Save Settings".

Scale Plotter Scope (manually)



In the four boxes surrounding this symbol you can enter the minimum x-coordinate, the maximum x-coordinate, the minimum y-coordinate and the maximum y-coordinate of the **Graph Drawer** as numbers.



Scale Plotter Scope (automatically)



Cause the **Graph Drawer** to set its scope to the data recently transmitted by clicking on this button.