

# How can I do the least squares problem in Maple?

## How can I do the least squares problem in Maple?

The method of least squares provides a means of estimating the values of coefficients in an equation. Typically, the estimates are based upon some sample of data. The idea of least squares is to minimize the total amount of error due to these estimates. Note that the technique of least squares is different from linear regression, though they have similar objectives.

Only discrete, finite least-squares applications will be discussed here. Two broad categories of such problems are:

- Equations that are linear in the coefficients. These can be solved using the Maple 'leastsqrs' function.

```
y = a + b x + c z
y = a x + b x^2 + c w^2 + d sin(x)
```

- Equations that are not linear in the coefficients. These cannot be solved using the Maple 'leastsqrs' function.

```
y = a + sin(b x)
y = (a x)^2 + log(b w)
```

The names {a,b,c,d} here are constants to estimate, and {w,x,y,z} are variables for which you would have sample data.

To use 'leastsqrs' for linear coefficients, do the following:

1. Arrange your equation in the form:

```
<dependent-variable> = <constant> + <constant><independent-term>
                        + <constant><independent-term> + ...
```

For example,  $y = a + b x$

2. Create a vector in Maple with all the "y" values for the data points. Lets assume the data you have is in the form (x,y) and consists of the points {(-1, 2), (0, 3), (1, 4.5), (2, 5)}.

```
> Y := array([2, 3, 4.5, 5]);
      Y := [ 2, 3, 4.5, 5 ]
```

3. Now create an array of the data for the independent variables, "x" in this particular example:

```
> A := array([ [1,-1], [1,0], [1,1], [1,2] ]);
      A := [
            [ 1 -1 ]
            [ 1  0 ]
            [ 1  1 ]
            [ 1  2 ]
            ]
```

Notice that the ones in the first column. This is because the desired equation has a constant term added, "a". If there were no such added constant term, this column would not be needed.

4. Now, load in the 'linalg' package to get the 'leastsqrs' function:

```
> with(linalg):
```

5. Now calculate the coefficients. They will be listed in the reverse order to the information in the data matrix, "A":

```
> leastsqrs(A,Y);
[ 3.100000000, 1.050000000 ]
```

So the estimated value of "b" is 3.1, and "a" is 1.05 in  $y = a + b x$ .

Here is an example when the equation is not linear (though the coefficients still are):

```
> # estimate coefficients for y = a x + b x^3
> # given data points (x,y) of {(0,0), (1,3), (2,18)}
> Y := array([0, 3, 18]);
      Y := [ 0, 3, 18 ]
> A := array([ [0,0], [1,1], [2,8] ]);
      A := [ 1 1 ]
           [ 2 8 ]
> leastsqrs(A,Y);
[ 1, 2 ]
```

Note that there was no initial column of ones, since there was no constant term being added to the equation.

As discussed above, you cannot use 'leastsqrs' for estimating non-linear use of constants. You can, if you want, perform a series expansion on a function and use this to "estimate" the estimates. Though simple in concept, the results may not be very accurate in practice. The function needs to be well behaved and the data set would need to be moderately large and well distributed.

The proper way to get the estimates for the constants in the non-linear case would be to derive an expression for the squared error and then use other Maple functions to minimize this expression. This is a standard problem in calculus and Maple has the tools to solve such problems. Assuming you only have one dependent variable (say, "y"), the steps you need to follow are:

1. Find the expression for the error at each point. For instance, consider the equation " $y = \sin(ax)$ " where "a" is the constant to estimate. The error in the estimate at a particular point (X,Y) would be " $Y - \sin(A X)$ " where "A" is denoting the estimated value of "a".
2. Square the error expression and sum it over all the data sets. For example, if you had (x,y) data points  $\{(0,0), (2,1)\}$  to fit the equation " $y = \sin(ax)$ " the Maple statement of the total squared error would be:

```
> errsqr := (0 - sin(a*0))^2 + (1 - sin(a*2))^2:
      errsqr := (1 - sin(2 a))^2
```

3. Take (partial) derivatives with respect to the constants to generate a system of equations and solve for the constants. By doing this you are finding the extremal points. You may need to check whether these extrema constitute minima or not by using the 2nd derivative test.

```
> solve( diff(errsqr,a) = 0, a );
      1/4 Pi, -1/4 Pi
> fsolve( diff(errsqr,a)=0, a );
      .7853962870
```

So the estimated value of the constant "a" is  $\pi/4$  or 0.7853962870. Note that you can only use the Maple 'solve' function when you are able to get a symbolic solution. 'fsolve' estimates the solutions numerically and can be applied to a wider range of non-linear equations.

You can find additional information by reading Maple's own on-line help on the following topics: leastsqrs, diff, solve, fsolve, array.

## See Also

- [Maple Landing Page](#)