

NDK_MLR_STEPWISE

Last Modified on 05/05/2016 12:28 pm CDT

- C/C++
- .Net

```
int __stdcall NDK_MLR_STEPWISE(double ** X,
                               size_t   nXSize,
                               size_t   nXVars,
                               LPBYTE   mask,
                               size_t   nMaskLen,
                               double *  Y,
                               size_t   nYSize,
                               double   intercept,
                               double   alpha,
                               WORD      nMode
                               )
```

Returns a list of the selected variables after performing the stepwise regression.

Returns

status code of the operation

Return values

NDK_SUCCESS Operation successful

NDK_FAILED Operation unsuccessful. See [Macros](#) for full list.

Parameters

- [in] **X** is the independent (explanatory) variables data matrix, such that each column represents one variable.
- [in] **nXSize** is the number of observations (rows) in X.
- [in] **nXVars** is the number of independent (explanatory) variables (columns) in X.
- [in,out] **mask** is the boolean array to choose the explanatory variables in the model. If missing, all variables in X are included.
- [in] **nMaskLen** is the number of elements in the "mask."
- [in] **Y** is the response or dependent variable data array (one dimensional array of cells).
- [in] **nYSize** is the number of observations in Y.
- [in] **intercept** is the constant or intercept value to fix (e.g. zero). If missing (i.e. NaN), an intercept will not be fixed and is computed normally.
- [in] **alpha** is the statistical significance of the test (i.e. alpha). If missing or omitted, an alpha value of 5% is assumed.
- [in] **nMode** is a switch to select the variable's inclusion/exclusion approach (1=forward selection (default), 2=backward elimination , 3=bi-directional elimination):
1. Forward selection

2. Backward elimination
3. Bi-directional elimination

Remarks

1. The underlying model is described [here](#).
2. The stepwise regression includes regression models in which the choice of predictive variables is carried out by an automatic procedure. The procedure takes the form of a sequence of f-tests in selecting or eliminating explanatory variables.
3. The three main approaches are:
 - **Forward Selection** which involves starting with no variables in the model, testing the addition of each variable using a chosen model comparison criterion, adding the variable (if any) that improves the model the most, and repeating this process until no additional variables improve the model.
 - **Backward Elimination** which involves starting with all candidate variables, testing the deletion of each variable using a chosen model comparison criterion, deleting the variable (if any) that improves the model the most by being deleted, and repeating this process until no further improvement is possible.
 - **Bidirectional Elimination** a combination of the above tests, involves testing at each step for variables to be included or excluded.
4. One of the main issues with stepwise regression is that it searches a large space of possible models. Hence it is prone to overfitting the data.
5. The initial values in the mask array define the variables set that MLR_STEPWISE works with. In other words, variables which are not selected will not be considered during the regression.
6. The sample data may include missing values.
7. Each column in the input matrix corresponds to a separate variable.
8. Each row in the input matrix corresponds to an observation.
9. Observations (i.e. row) with missing values in X or Y are removed.
10. The number of rows of the response variable (Y) must be equal to the number of rows of the explanatory variables (X).
11. The MLR_STEPWISE function is available starting with version 1.60 APACHE.

Requirements

Header	SFSDK.H
Library	SFSDK.LIB
DLL	SFSDK.DLL

```

    UIntPtr    nXSize,
    UIntPtr    nXVars,
    byte[]     mask1,
    byte[]     mask2,
    UIntPtr    nMaskLen,
    double[]   pYData,
    UIntPtr    nYSize,
    double     intercept,
    double     alpha,
    short      nRetType,
    ref double retVal
)

```

Class: SFSDK
Scope: Public
Lifetime: Static

Returns a list of the selected variables after performing the stepwise regression.

Returns

status code of the operation

Return values

NDK_SUCCESS Operation successful

NDK_FAILED Operation unsuccessful. See [Macros](#) for full list.

Parameters

- [in] **pXData** is the independent (explanatory) variables data matrix, such that each column represents one variable.
- [in] **nXSize** is the number of observations (rows) in pXData.
- [in] **nXVars** is the number of independent (explanatory) variables (columns) in pXData.
- [in,out] **mask** is the boolean array to choose the explanatory variables in the model. If missing, all variables in pXData are included.
- [in] **nMaskLen** is the number of elements in the "mask."
- [in] **pYData** is the response or dependent variable data array (one dimensional array of cells).
- [in] **nYSize** is the number of observations in pYData.
- [in] **intercept** is the constant or intercept value to fix (e.g. zero). If missing (i.e. NaN), an intercept will not be fixed and is computed normally.
- [in] **alpha** is the statistical significance of the test (i.e. alpha). If missing or omitted, an alpha value of 5% is assumed.
- [in] **nMode** is a switch to select the variable's inclusion/exclusion approach (1=forward selection (default), 2=backward elimination , 3=bi-directional elimination):
 1. Forward selection
 2. Bacward elimination
 3. Bi-directional elemination

Remarks

1. The underlying model is described [here](#).
2. The stepwise regression includes regression models in which the choice of predictive variables is carried out by an automatic procedure. The procedure takes the form of a sequence of f-tests in selecting or eliminating explanatory variables.
3. The three main approaches are:
 - **Forward Selection** which involves starting with no variables in the model, testing the addition of each variable using a chosen model comparison criterion, adding the variable (if any) that improves the model the most, and repeating this process until no additional variables improve the model.
 - **Backward Elimination** which involves starting with all candidate variables, testing the deletion of each variable using a chosen model comparison criterion, deleting the variable (if any) that improves the model the most by being deleted, and repeating this process until no further improvement is possible.
 - **Bidirectional Elimination** a combination of the above tests, involves testing at each step for variables to be included or excluded.
4. One of the main issues with stepwise regression is that it searches a large space of possible models. Hence it is prone to overfitting the data.
5. The initial values in the mask array define the variables set that MLR_STEPWISE works with. In other words, variables which are not selected will not be considered during the regression.
6. The sample data may include missing values.
7. Each column in the input matrix corresponds to a separate variable.
8. Each row in the input matrix corresponds to an observation.
9. Observations (i.e. row) with missing values in X or Y are removed.
10. The number of rows of the response variable (Y) must be equal to the number of rows of the explanatory variables (X).
11. The MLR_STEPWISE function is available starting with version 1.60 APACHE.

Exceptions

Exception Type	Condition
None	N/A

Requirements

Namespace	NumXLAPI
Class	SFSDK
Scope	Public
Lifetime	Static

Package	NumXLAPI.DLL
----------------	--------------

Examples

References

Hamilton, J .D.; [Time Series Analysis](#) , Princeton University Press (1994), ISBN 0-691-04289-6

Tsay, Ruey S.; [Analysis of Financial Time Series](#) John Wiley & SONS. (2005), ISBN 0-471-690740

See Also

[template("related")]
