R FUNCTIONS FOR REGRESSION ANALYSIS

Here are some helpful R functions for regression analysis grouped by their goal. The name of package is in parentheses.

Linear model

Anova: Anova Tables for Linear and Generalized Linear Models (car)
anova: Compute an analysis of variance table for one or more linear model fits (stasts)
coef: is a generic function which extracts model coefficients from objects returned by modeling functions. coefficients is an alias for it (stasts)
coeftest: Testing Estimated Coefficients (lmtest)
confint: Computes confidence intervals for one or more parameters in a fitted model. Base has a method for objects inheriting from class "lm" (stasts)
deviance:Returns the deviance of a fitted model object (stats)
effects: Returns (orthogonal) effects from a fitted model, usually a linear model. This is a generic function, but currently only has a methods for objects inheriting from classes "lm" and "glm" (stasts)
fitted: is a generic function which extracts fitted values from objects returned by modeling functions fitted.values is an alias for it (stasts)
formula: provide a way of extracting formulae which have been included in other objects (stasts)
linear.hypothesis: Test Linear Hypothesis (car)
lm: is used to fit linear models. It can be used to carry out regression, single stratum analysis of variance and analysis of covariance (stasts)
model.matrix: creates a design matrix (stasts)
predict: Predicted values based on linear model object (stasts)
residuals: is a generic function which extracts model residuals from objects returned by modeling functions (stasts)
summary.lm: summary method for class "lm" (stats)
vcov: Returns the variance-covariance matrix of the main parameters of a fitted model object (stasts)

Model – Variables selection

add1: Compute all the single terms in the scope argument that can be added to or dropped from the model, fit those models and compute a table of the changes in fit (stats)
AIC: Generic function calculating the Akaike information criterion for one or several fitted model objects for which a log-likelihood value can be obtained, according to the formula $-2*log-likelihood + k*npar$, where npar represents the number of parameters in the fitted model, and k = 2 for the usual AIC, or k = log(n) (n the number of observations) for the so-called BIC or SBC (Schwarz's Bayesian criterion) (stats)
Cpplot: Cp plot (faraway)
drop1: Compute all the single terms in the scope argument that can be added to or dropped from the model, fit those models and compute a table of the changes in fit (stats)
extractAIC: Computes the (generalized) Akaike An Information Criterion for a fitted parametric model (stats)
leaps: Subset selection by `leaps and bounds' (leaps)
maxadjr: Maximum Adjusted R-squared (faraway)
offset: An offset is a term to be added to a linear predictor, such as in a generalised linear model, with known coefficient 1 rather than an estimated coefficient (stats)
step: Select a formula-based model by AIC (stats)
update.formula: is used to update model formulae. This typically involves adding or dropping terms, but updates can be more general (stats)

Diagnostics
cookd: Cook's Distances for Linear and Generalized Linear Models (car)
cooks.distance: Cook's distance (stats)
covratio: covariance ratio (stats)
dfbeta: DBETA (stats)
dfbetas: DBETAS (stats)
dffits: DFFTITS (stats)
hat: diagonal elements of the hat matrix (stats)
hatvalues: diagonal elements of the hat matrix (stats)
influence.measures: This suite of functions can be used to compute some of the regression (leave-one-out deletion) diagnostics for linear and generalized linear models (stats)
lm.influence: This function provides the basic quantities which are used in forming a wide variety of diagnostics for checking the quality of regression fits (stats)
ls.diag: Computes basic statistics, including standard errors, t- and p-values for the regression coefficients (stats)
outlier.test: Bonferroni Outlier Test (car)
rstandard: standardized residuals (stats)
rstudent: studentized residuals (stats)
vif: Variance Inflation Factor (car)

Graphics
ceres.plots: Ceres Plots (car)
cr.plots: Component+Residual (Partial Residual) Plots (car)
influence.plot: Regression Influence Plot (car)
leverage.plots: Regression Leverage Plots (car)
panel.car: Panel Function Coplots (car)
plot.lm: Four plots (selectable by which) are currently provided: a plot of residuals against fitted values, a Scale-Location plot of sqrt{| residuals |} against fitted values, a Normal Q-Q plot, and a plot of Cook's distances versus row labels (stats)
prplot: Partial Residual Plot (faraway)
qq.plot: Quantile-Comparison Plots (car)
qqline: adds a line to a normal quantile-quantile plot which passes through the first and third quartiles (stats)
qqnorm: is a generic function the default method of which produces a normal QQ plot of the values in y (stats)
reg.line: Plot Regression Line (car)
scatterplot.matrix: Scatterplot Matrices (car)
scatterplot: Scatterplots with Boxplots (car)
spread.level.plot: Spread-Level Plots (car)

Tests
ad.test: Anderson-Darling test for normality (nortest)
bartlett.test: Performs Bartlett's test of the null that the variances in each of the groups (samples) are the same (stats)
bgtest: Breusch-Godfrey Test (lmtest)
bptest: Breusch-Pagan Test (lmtest)
cvm.test: Cramer-von Mises test for normality (nortest)
durbin.watson: Durbin-Watson Test for Autocorrelated Errors (car)
dwtest: Durbin-Watson Test (lmtest)
levene.test: Levene's Test (car)
lillie.test: Lilliefors (Kolmogorov-Smirnov) test for normality (nortest)
ncv.test: Score Test for Non-Constant Error Variance (car)
pearson.test: Pearson chi-square test for normality (nortest)
sf.test: Shapiro-Francia test for normality (nortest)

shapiro.test: Performs the Shapiro–Wilk test of normality (stats)

Variables transformations
box.cox: Box-Cox Family of Transformations (car)
boxcox: Box-Cox Transformations for Linear Models (MASS)
box.cox.powers: Multivariate Unconditional Box-Cox Transformations (car)
box.tidwell: Box-Tidwell Transformations (car)
box.cox.var: Constructed Variable for Box-Cox Transformation (car)

Ridge regression
lm.ridge: Ridge Regression (MASS)

Segmented regression
segmented: Segmented relationships in regression models (segmented)
slope.segmented: Summary for slopes of segmented relationships (segmented)

Generalized Least Squares (GLS)
ACF.gls: Autocorrelation Function for gls Residuals (nlme)
anova.gls: Compare Likelihoods of Fitted Objects (nlme)
gls: Fit Linear Model Using Generalized Least Squares (nlme)
intervals.gls: Confidence Intervals on gls Parameters (nlme)
lm.gls: fit Linear Models by Generalized Least Squares (MASS)
plot.gls: Plot a gls Object (nlme)
predict.gls: Predictions from a gls Object (nlme)
qqnorm.gls: Normal Plot of Residuals from a gls Object (nlme)
residuals.gls: Extract gls Residuals (nlme)
summary.gls: Summarize a gls Object (nlme)

Generalized Linear Models (GLM)
family: Family objects provide a convenient way to specify the details of the models used by functions such as glm (stats)
glm.nb: fit a Negative Binomial Generalized Linear Model (MASS)
glm: is used to fit generalized linear models, specified by giving a symbolic description of the linear predictor and a description of the error distribution (stats)
polr: Proportional Odds Logistic Regression (MASS)

Non linear Least Squares (NLS)
nlm: This function carries out a minimization of the function f using a Newton-type algorithm (stats)
nls: Determine the nonlinear least-squares estimates of the nonlinear model parameters and return a class nls object (stats)
nlscontrol: Allow the user to set some characteristics of the nls nonlinear least squares algorithm (stats)
nlsModel: This is the constructor for nlsModel objects, which are function closures for several functions in a list. The closure includes a nonlinear model formula, data values for the formula, as well as parameters and their values (stats)

Generalized Non linear Least Squares (GNLS)
coef.gnls: Extract gnls Coefficients (nlme)
gnls: Fit Nonlinear Model Using Generalized Least Squares (nlme)
predict.gnls: Predictions from a gnls Object (nlme)

Loess regression

loess: Fit a polynomial surface determined by one or more numerical predictors, using local fitting (stats)

loess.control: Set control parameters for loess fits (stats)

predict.loess: Predictions from a loess fit, optionally with standard errors (stats)

scatter.smooth: Plot and add a smooth curve computed by loess to a scatter plot (stats)

Splines regression

bs: B-Spline Basis for Polynomial Splines (splines)
ns: Generate a Basis Matrix for Natural Cubic Splines (splines)
periodicSpline: Create a Periodic Interpolation Spline (splines)
polySpline: Piecewise Polynomial Spline Representation (splines)
predict.bSpline: Evaluate a Spline at New Values of x (splines)
predict.bs: Evaluate a Spline Basis (splines)
splineDesign: Design Matrix for B-splines (splines)
splineKnots: Knot Vector from a Spline (splines)
splineOrder: Determine the Order of a Spline (splines)

Robust regression

lqs: Resistant Regression (MASS)
rlm: Robust Fitting of Linear Models (MASS)

Structural equation models

sem: General Structural Equation Models (sem)
tsls: Two-Stage Least Squares (sem)

Simultaneous Equation Estimation

systemfit: Fits a set of linear structural equations using Ordinary Least Squares (OLS), Weighted Least Squares (WLS), Seemingly Unrelated Regression (SUR), Two-Stage Least Squares (2SLS), Weighted Two-Stage Least Squares (W2SLS) or Three-Stage Least Squares (3SLS) (systemfit)

Partial Least Squares Regression (PLSR) and Principal Component Regression (PCR)

biplot.mvr: Biplots of PLSR and PCR Models (pls)
coefplot: Plot Regression Coefficients of PLSR and PCR models (pls)
crossval: Cross-validation of PLSR and PCR models (pls)
cvsegments: Generate segments for cross-validation (pls)
kernelpls.fit: Kernel PLS (Dayal and MacGregor) (pls)
msc: Multiplicative Scatter Correction (pls)
mvr: Partial Least Squares and Principal Components Regression (pls)
mvrCv: Cross-validation (pls)
oscorespls.fit: Orthogonal scores PLSR (pls)
predplot: Prediction Plots (pls)
scoreplot: Plots of Scores and Loadings (pls)
scores: Extract Scores and Loadings from PLSR and PCR Models (pls)
svdpc.fit: Principal Components Regression (pls)
validationplot: Validation Plots (pls)

Quantile regression

anova.rq: Anova function for quantile regression fits (quantreg)
boot.rq: Bootstrapping Quantile Regression (quantreg)
lprq: locally polynomial quantile regression (quantreg)

nlrq: Function to compute nonlinear quantile regression estimates (quantreg)
qss: Additive Nonparametric Terms for rqss Fitting (quantreg)
ranks: Quantile Regression Ranks (quantreg)
rq: Quantile Regression (quantreg)
rqss: Additive Quantile Regression Smoothing (quantreg)
rrs.test: Quantile Regression Rankscore Test (quantreg)
standardize: Function to standardize the quantile regression process (quantreg)

Linear and nonlinear mixed effects models
ACF: Autocorrelation Function (nlme)
ACF.lme: Autocorrelation Function for lme Residuals (nlme)
anova.lme: compare Likelihoods of Fitted Objects (nlme)
fitted.lme: Extract lme Fitted Values (nlme)
fixed.effects: Extract Fixed Effects (nlme)
intervals: Confidence Intervals on Coefficients (nlme)
intervals.lme: Confidence Intervals on lme Parameters (nlme)
lme: Linear Mixed-Effects Models (nlme)
nlme: Nonlinear Mixed-Effects Models (nlme)
predict.lme: Predictions from an lme Object (nlme)
predict.nlme: Predictions from an nlme Obj (nlme)
qqnorm.lme: Normal Plot of Residuals or Random Effects from an lme object (nlme)
random.effects: Extract Random Effects (nlme)
ranef.lme: Extract lme Random Effects (nlme)
residuals.lme: Extract lme Residuals (nlme)
simulate.lme: simulate lme models (nlme)
summary.lme: Summarize an lme Object (nlme)
glmmPQL: fit Generalized Linear Mixed Models via PQL (MASS)

Generalized Additive Model (GAM)
anova.gam: compare the fits of a number of gam models (gam)
gam.control: control parameters for fitting gam models (gam)
gam: Fit a generalized additive model (gam)
na.gam.replace: a missing value method that is helpful with gams (gam)
plot.gam: an interactive plotting function for gams (gam)
predict.gam: make predictions from a gam object (gam)
preplot.gam: extracts the components from a gam in a plot-ready form (gam)
step.gam: stepwise model search with gam (gam)
summary.gam: summary method for gam (gam)

Survival analysis
anova.survreg: ANOVA tables for survreg objects (survival)
clogit: Conditional logistic regression (survival)
cox.zph: Test the proportional hazards assumption of a Cox regression (survival)
coxph: Proportional Hazards Regression (survival)
coxph.detail: Details of a cox model fit (survival)
coxph.rvar: Robust variance for a Cox model (survival)
ridge: ridge regression (survival)
survdiff: Test Survival Curve Differences (survival)
survexp: Compute Expected Survival (survival)
survfit: Compute a survival Curve for Censored Data (survival)
survreg: Regression for a parametric survival model (survival)

Classification and Regression Trees
cv.tree: Cross-validation for Choosing tree Complexity (tree)
deviance.tree: Extract Deviance from a tree Object (tree)
labels.rpart: Create Split Labels For an rpart Object (rpart)
meanvar.rpart: Mean-Variance Plot for an rpart Object (rpart)

misclass.tree: Misclassifications by a Classification tree (tree)
na.rpart: Handles Missing Values in an rpart Object (rpart)
partition.tree: Plot the Partitions of a simple Tree Model (tree)
path.rpart: Follow Paths to Selected Nodes of an rpart Object (rpart)
plotcp: Plot a Complexity Parameter Table for an rpart Fit (rpart)
printcp: Displays CP table for Fitted rpart Object (rpart)
prune.misclass: Cost-complexity Pruning of Tree by error rate (tree)
prune.rpart: Cost-complexity Pruning of an rpart Object (rpart)
prune.tree: Cost-complexity Pruning of tree Object (tree)
rpart: Recursive Partitioning and Regression Trees (rpart)
rpconvert: Update an rpart object (rpart)
rsq.rpart: Plots the Approximate R–Square for the Different Splits (rpart)
snip.rpart: Snip Subtrees of an rpart Object (rpart)
solder: Soldering of Components on Printed–Circuit Boards (rpart)
text.tree: Annotate a Tree Plot (tree)
tile.tree: Add Class Barplots to a Classification Tree Plot (tree)
tree.control: Select Parameters for Tree (tree)
tree.screens: Split Screen for Plotting Trees (tree)
tree: Fit a Classification or Regression Tree (tree)

Beta regression
betareg: Fitting beta regression models (betareg)
plot.betareg: Plot Diagnostics for a betareg Object (betareg)
predict.betareg: Predicted values from beta regression model (betareg)
residuals.betareg: Residuals function for beta regression models (betareg)
summary.betareg: Summary method for Beta Regression (betareg)