

Appendix E

MATLAB Functions

For the convenience of the reader, we provide a list of functions and commands that were mentioned in the text and others that we think might be of interest to the exploratory data analyst. Be sure to check the **help** files on the Statistics Toolbox and EDA Toolbox for a list of all functions that are available.

E.1 MATLAB

abs	Absolute value of a vector or matrix.
axes	Low-level function to set axes properties.
axis	Command or function to change axes.
bar, bar3	Construct a 2-D or 3-D bar graph.
cart2pol	Change from Cartesian to polar coordinates.
ceil	Round up to the next integer.
char	Creates character array.
cla	Clears the current axes.
colorbar	Puts a color scale on the figure.
colormap	Changes the color map for the figure.
contour	Constructs a contour plot.
convhull	Finds the convex hull.
cos	Finds the cosine of the angle (radians).
cov	Returns the sample covariance matrix.
cumsum	Calculates the cumulative sum of vector elements.
diag	Diagonal matrices and diagonals of a matrix.
diff	Finds the difference between elements.
drawnow	Force MATLAB to update the graphics screen.
eig, eigs	Eigenvalues/eigenvectors for full and sparse matrices.
exp	Find the exponential of the argument.
eye	Provides an identity matrix.
figure	Creates a blank figure window.
find	Find indices of nonzero elements.
findobj	Find objects with certain properties.
flipud	Flip matrix in up/down direction.

floor	Round down to the nearest integer.
gammaln	Logarithm of the gamma function.
gray	A pre-specified gray-scale color map.
help <i>funcname</i>	Command line way to get help on a function.
hist	Create a histogram.
hold	Freeze the current plot, graphics are added.
imagesc	Scaled image of the data.
inpolygon	True if the point is inside the polygon, zero otherwise.
int2str	Convert integer to char.
intersect	Elements in the set intersection.
legend	Add a legend to the axes.
length	Returns the length of a vector.
line	Low-level function to create a line or curve.
linspace	Provide linearly-spaced points in a given 1-D range.
load	Load some data – MAT-file or ASCII.
log	Take logarithm of elements.
mean	Find the sample mean.
median	Find the median.
meshgrid	Provide linearly-spaced points (X, Y) in a 2-D domain.
min, max	Find the minimum or maximum element in a vector.
mod	Returns modulus after division.
norm	Calculates the vector or matrix norm.
ones	Provides an array with all ones.
pause	Pause all activity in MATLAB.
pcolor	Pseudo-color plot – elements specify colors.
pinv	Pseudo-inverse of a matrix.
plot, plot3	2-D or 3-D plots.
plotmatrix	Matrix of scatterplots.
pol2cart	Convert polar to Cartesian coordinates.
polyfit	Fit data to a polynomial.
polyval	Evaluate a polynomial.
rand	Generate uniform (0, 1) random variables.
randn	Generate standard normal random variables.
repmat	Construct a new array with replicated/tiled elements.
reshape	Reshape or change the size of a matrix.
round	Round towards the nearest integer.
scatter, scatter3	Construct 2-D and 3-D scatterplots.
semilogy	Construct a 2-D plot, with a vertical log scale.
set	Set properties of graphics objects.
setdiff	Find the set difference between vectors.
sin	Calculate the sine (radians.)
size	Find the size of an array.
sort	Sort in ascending order.
sqrt	Find the square root.
std	Find the sample standard deviation.
strmatch	Find matches for a string.
sum	Add up the elements in a vector.
surf	Construct a surface plot.
svd	Singular value decomposition.
tan	Find the tangent of an angle (radians).

title	Put a title on a plot.
vander	Vandermonde matrix.
xlabel, ylabel	Add labels to the axes in a plot.
zeros	Returns an array with all zeros.

E.2 Statistics Toolbox - Versions 4 and 5

Unless otherwise noted, the following functions are included in version 4 of the Statistics Toolbox.

biplot (V5)	Biplot of variable/factor coefficients and scores.
boxplot	Displays boxplots of the columns of a matrix.
cluster	Get clusters from hierarchical clustering.
clusterdata	Construct clusters from data (whole process).
cmdscale	Classical multi-dimensional scaling.
cophenet	Calculates the cophenetic coefficient.
dendrogram	Constructs the dendrogram.
dfittool (V5)	GUI for fitting distributions to data.
exprnd	Generate random variables from exponential.
factoran	Factor analysis.
gname	Interactive labeling of scatterplot.
gplotmatrix	Grouped scatterplot matrix.
gscatter	Grouped scatter plot.
hist3 (V5)	Histogram of bivariate data.
iqr	Interquartile range.
kmeans	k-means clustering.
linkage	Agglomerative hierarchical clustering.
mdscale (V5)	Metric and nonmetric multidimensional scaling.
mvnpdf	Multivariate normal probability density function.
mvnrnd	Generate multivariate normal random variables.
norminv	Inverse of the normal cumulative distribution function.
normpdf	1-D normal probability density function.
normplot	Normal probability plot.
normrnd	Generate 1-D normal random variables.
pcacov	PCA using the covariance matrix.
pdist	Find interpoint distances.
probplot (V5)	Probability plots for specified distributions.
princomp	Principal component analysis.
qqplot	Empirical q-q plot
quantile (V5)	Find the quantiles of a sample.
silhouette	Construct silhouette plot and return silhouette values.
squareform	Convert output of pdist to matrix.
unifrnd	Generate uniform random variables over range (a, b).
weibplot	Weibull probability plot.

E.3 Exploratory Data Analysis Toolbox

adjrand	Adjusted Rand index to compare groupings.
agmclust	Model-based agglomerative clustering.
bagmat.exe	Executable file to get arrays needed for bagplot.
bagplot	M-file to construct actual bagplot.
boxp	Boxplot - regular.
boxprct	Box-percentile plot.
brushscatter	Scatterplot brushing and linking.
coplot	Coplot from Data Visualization Toolbox.
csandrews	Andrews' curves plot.
csparallel	Parallel coordinates plot.
csppstrtrem	Remove structure in PPEDA.
dotchart	Dot chart plot.
genmix	GUI to generate random variables from finite mixture.
gtm_pmd	Calculates posterior mode projection (GTM Toolbox).
gtm_pmn	Calculates posterior mean projection (GTM Toolbox).
gtm_stp2	Generates components of a GTM (GTM Toolbox).
gtm_trn	Train the GTM using EM (GTM Toolbox).
hexplot	Hexagonal binning for scatterplot.
hlle	Hessian eigenmaps.
idpettis	Intrinsic dimensionality estimate.
intour	Interpolation tour of the data.
isomap	ISOMAP nonlinear dimensionality reduction.
kdimtour	k -dimensional grand tour.
lle	Locally linear embedding.
loess	1-D loess scatterplot smoothing.
loess2	2-D loess smoothing from Data Visualization Toolbox.
loessenv	Loess upper and lower envelopes.
loessr	Robust loess scatterplot smoothing.
mbcfmix	Model-based finite mixture estimation - EM.
mbclust	Model-based clustering.
mixclass	Classification using mixture model.
multiwayplot	Multiway dot charts.
nmmds	Nonmetric multidimensional scaling.
permtourandrews	Permutation tour using Andrews' curves.
permtourparallel	Permutation tour using parallel coordinate plots.
plotbic	Plot the BIC values from model-based clustering.
plotmatrixandr	Plot matrix of Andrews' curves.
plotmatrixpara	Plot matrix of parallel coordinates.
polarloess	Bivariate smoothing using loess.
ppeda	Projection pursuit EDA.
pseudotour	Pseudo grand tour.
quantileseda	Sample quantiles.
quartiles	Sample quartiles using Tukey's fourths.
randind	Rand index to compare groupings.
reclus	ReClus plot to visualize cluster output.

rectplot	Rectangle plot to visualize hierarchical clustering.
scattergui	Scatterplot with interactive labeling.
som_autolabel	Automatic labeling (SOM Toolbox).
som_data_struct	Create a data structure (SOM Toolbox).
som_make	Create, initialize and train SOM (SOM Toolbox).
som_normalize	Normalize data (SOM Toolbox).
som_set	Set up SOM structures (SOM Toolbox).
som_show	Basic SOM visualization (SOM Toolbox).
som_show_add	Shows hits, labels and trajectories (SOM Toolbox).
torustour	Asimov grand tour
treemap	Treemap display for hierarchical clustering.