

# Visual Data Analysis Tools for Scientists and Engineers

Willy Hereman

## 1. Introduction

In this article a plethora of mathematics and graphics packages are reviewed. MATLAB, O-MATRIX, GAUSS, Speakeasy and Mathcad, all target the market of researchers who use matrix-based modeling and analysis, and visualization tools on a PC. As of this year, the computer algebra package MathPlus is available for PC Windows. Although this product is different in scope, its fascinating features persuaded me to include it in this survey. This article will highlight these six software packages and how they address different user needs, have unique features, capabilities, various degrees of sophistication in programming environments, and how they cover a wide variety of applications.

I used an IBM compatible 486 PC, with DX2-66 MHz processor, 16 MB of RAM, 425 MB hard disk, and a 4 MB swap file. The PC has MicroSoft Windows v. 3.1 and DOS 6.0, and is supported by a Hewlett Packard LaserJet III (without PostScript) and a Super VGA color monitor. If you want to produce paper copies of graphs produced by the above toolboxes, I recommend using a laser printer with PostScript.

The installation of all six packages was straightforward and flawless. Each package was tested separately, one copy of each program at the time, and without other background applications. Under Windows, multi-tasking is possible: graphics, commands and text can be handled in separated windows, pull-down menus and hypertext links facilitate fast access to on-line help. A click on highlighted text on the help calls up help screens with related information. The efficiency of these searches makes thumbing through printed manuals almost unnecessary. The packages MATLAB, O-MATRIX, MathPlus, and Mathcad make use of these navigation tools and other menu-based facilities under Windows. Speakeasy has its own, quite elegant, window system. GAUSS does not have window capabilities, but info is provided via help screens.

## 2. MathPlus

Theorist, a symbolic algebra and graphics package for Macintosh, is now available for Windows as MathPlus. MathPlus is a mouse-oriented computer algebra system, not as comprehensive in scope and functionality as the brothers Mathematica, Macsyma, and Maple, but much easier to use due to its unique graphical user interface. MathPlus is one of the few programs to take a step towards natural, transparent and direct manipulation of mathematical objects.

Working with MathPlus is fun. In an electronic notebook, you create mathematical

expressions by picking them from a palette with point-and-click icons, or by entering them via the keyboard. The equations on screen can be edited with a built-in WYSIWYG editor. Using a 'click and solve' technology, expressions and terms can be moved around the screen. People familiar with the symbolic package Milo, the math editor Expressionist, or Mathcad will feel right at home with MathPlus. MathPlus mimics the way humans do mathematics on paper. Unlike its large competitors, it does not require you to learn a specific syntax or programming language to do it. Its elegance of use reminded me of working with DERIVE, a symbolic manipulator for PC under DOS, reviewed by John Fitch in the June 1993 issue of this magazine.

With MathPlus you can solve college level problems numerically or symbolically, demonstrate mathematical concepts, and visualize solutions to problems via 2 and 3D graphs and QuickTime movies. It is an excellent assistant for students or teachers who want to explore mathematical concepts or solve problems from mathematics, physics or engineering. The 'super-calculator' MathPlus performs algebra, trigonometry, matrix algebra, standard operations from calculus, vector calculus and statistics. Nearly 300 math functions are built in, including many special functions. The error function, the Bessel, Gamma, Beta, and the Riemann  $\zeta$ -function are all implemented. In addition to a wide variety of pre-defined functions, MathPlus allows users to create new functions, and work with external data. Table functions can operate as alternatives to formulas and graphs for representing real and complex functions.

Instructive live notebooks illustrate finite and infinite series, ideas from number theory, Laplace transforms, graphics in 3-D, units, and more. The program supports non-commutative and non-associative algebra, and allows one to explore permutation groups. It can solve simple differential equations analytically or numerically, and it allows one to compute Fourier and Laplace transforms. The FFT routine on external data sets is limited, however, to 1024 points. However, due to its limitations in nonlinear mathematics, fairly primitive calculus routines, and interactive nature, I would not recommend MathPlus for advanced math projects.

MathPlus handles graphics with exemplary simplicity. The graphical capabilities include 2-D and 3-D graphics in cartesian and curvilinear coordinates, with various forms of animation. A unique feature allows you to store real or complex values in a Table, which can be displayed graphically. A drawback is that the Table function only stores the function values, assuming that the values are given at regular, equally spaced intervals.

On the negative side, MathPlus does not offer full programmability, although the user can set up rules for pattern matching. Few add-ons for engineering applications are available, and the limitations on the size of problems with integers and matrices are not clear from the documentation. On the positive side, the program is easy to figure out by yourself. The learning guide, the reference manual, the on-line help (with hypertext search facilities) and the live notebooks are excellent.

### 3. O-MATRIX

As a new addition to mathematical software by Harmonic Software Inc., O-MATRIX is an interactive data-analysis package and visualization tool. Built-in subroutines provide analysis power based on matrices operations such as singular value decomposition, QR and Cholesky factorization, matrix exponentiation, solution of Toeplitz systems, etc. The functions within the O-matrix environment are optimized for speed which makes the development of fast-executing applications easy. O-MATRIX uses the Rational System's DOS Extender so that programs and data arrays can use up to 32 MB of memory.

O-MATRIX does more than remove the drudgery of numerical computations based on matrix algebra. It also adds various routines for solving ODEs, e.g. via the 4th-order Runge-Kutta method. The major application of O-MATRIX is object-oriented visual data analysis. It has a diverse toolbox for numerical analysis, including programs for one and 2D discrete FFTs, fitting with cubic splines, optimization techniques, and linear and nonlinear least squares routines. The package supports Kalman, Butterworth and Chebyshev filters, tools for spectral estimation, statistics, and ARIMA simulation.

You can create a variety of plots, including surface and contour plots. O-MATRIX provides interactive graphics, where the user can, for example, rotate graphs, specify the labels on axes, and change the aspect ratio. Unfortunately, at this point no animation is provided.

Keep in mind that O-MATRIX works with vectors and matrices, which make its graphics less transparent than in computer algebra packages. O-MATRIX is picky about data types, similar to Fortran and C. Asking for powers of matrices with integer entries caused trouble, presumably due to the fact that O-MATRIX allocates a fixed amount of memory to mathematical objects. If an integer becomes too large to fit as an exact number, the program displays an error message.

O-matrix comes with a built-in-editor, a debugger that helps trouble-shoot, and a transparent on-line help system, which offers efficient interactive hierarchical searches. The program is easy to use, if you are willing to learn its specific syntax, which resembles Pascal. As with MATLAB, you can write your own custom functions and 'include' them in the program. Several demonstration files are available and most of the code can be inspected with your favorite editor. In a couple of hours you get a good idea of what O-MATRIX can and cannot do. The examples in the User's Guide are too simple and not very informative. At the time of this review, the Function Reference Guide, which highlights the functions in the 'toolbox', was not yet available.

The attractive price of O-MATRIX and its ease of use make it a nice alternative to MATLAB or GAUSS, but you receive a less comprehensive function list and more rudimentary plotting capabilities.

## 4. Speakeasy

Speakeasy is a DOS-based, fairly small, yet powerful computing environment for mathematical and statistical modeling and data analysis. Although it does not run under MicroSoft Windows, the designers have developed an impressive DOSWINDOWS system, that supports a mouse, and uses highlighting techniques for searches in the on-line documentation (similar to hypertext). I tested two versions of the software: the Eta version and the Beta test release of Eta+. In the Eta+ release, several Speakeasy commands have been modified for better performance, a NEWS help document has been added, and functions are better documented. Among the new features, highlighting of matching parentheses and multiple windows for graphics output are wonderful.

Although Speakeasy uses a line-oriented editor, you can switch to your favorite screen editor. Therefore, Speakeasy offers interactive and non-interactive environments. The user can enter a request and see the results immediately, or, for more complicated projects with larger data, and work with batch files.

The syntax of Speakeasy looks like a simplified version of Fortran, but is such that commands can be entered with a minimal number of keystrokes. You can recall up to two dozen previously entered lines, define keys to be the equivalent of typed input, and cut and paste with the mouse.

After running the built-in overview and playing with the help system, I was convinced that Speakeasy has much more power than one might expect at first glance. You can choose from differential equations, FFT transforms, data fitting, LINPACK routines for vector and matrix manipulations, optimization, linear programming, statistical analysis, financial mathematics, set algebra, time series analysis, physics, numerical differentiation and integration, graphics, etc.. The graphics include curves, 3-D graphs, contour plots, and business graphs, such as pie and bar charts.

Among the 750 mathematical functions, you'll find differential equations solvers based on the Adam-Moulton and Shampine methods, the Kolmogorov-Smirnov test in non-parametric statistics, routines for the Clebsch-Gordon and Wigner coefficients in physics, various random number generators, routines to process characters, to name a few of the more exotic examples. There is even a 'whoops' function that lets you cancel an erroneously entered modification.

Speakeasy has complete on-line documentation for all of its facilities. The Help and Document menus offer quick and logical access to information, and the user can create additional menus. The printed manuals that come with Speakeasy are exemplify organization and clarity. I was impressed with the 'Getting Started' and the 'Vocabulary Overview' 80 page booklets. Rarely seen, in a handy two-way dictionary, commands are alphabetically ordered by name as well as grouped by functionality. I wish there was a concise booklet showing a dozen worked applications of Speakeasy, though samples of interactive sessions in the manual are a step in the right direction.

The Speakeasy philosophy is that the user should concentrate on the specific science problem, rather than worry about the computer. In that respect, the program is quite unique:

commands are similar to traditional mathematical notation, commonly used operations are pre-programmed, a simple debugger is available. The printed documentation is concise and accurate, the on-line documentation and its access are excellent.

Speakeasy is a wonderful tool. It teams ease-of-use with advanced features, making it a good choice for novices as well as experienced scientists. If Speakeasy keeps growing it could well become one of the best packages available for data analysis and numerical computations.

## 5. Mathcad

Mathcad, under development by MathSoft since 1985, remains a product of choice for those who want an easy-to-use numerical and symbolic package. Mathcad PLUS 5.0 is the most advanced version of Mathcad. It is designed for the mainstream technical professional who needs advanced math functionality for data analysis and visualization.

The PLUS version incorporates all of the new features and usability enhancements of Mathcad 5.0 such as improved equation and document editing, animated tutorials, a spell checker with math dictionary, a print previewer, tracing of data and coordinates, zooming of plots, etc.. The PLUS version includes a dozen differential equation solvers, advanced matrix functions such as SVD, Cholesky square root, and generalized eigenvalues, more symbolic and numeric calculation power, and a 32-bit C and C++ programming interface.

Just like MathPlus, Mathcad's WYSISWG interface lets you see formulas instantly, in real math notation: two story fractions, integral signs, matrices with brackets around them, summation and product symbols. With Mathcad's Live Document Interface you can mix readable formulas, tables, attractive graphics and commentary, thus creating a live document on screen or a professional-looking printed report. Indeed, the equations and graphs are live. Any changes to data or equations are automatically reflected in numeric or symbolic results or graphs. This integrated environment— an electronic whiteboard— is one of the nicest features of Mathcad.

In Mathcad it is easy to enter mathematics, with subscripts and superscripts, but you have to remember a few awkward keystrokes like '[ctrl][Shift]4' for a summation symbol, '?' for the derivative, and '&' for an integral. The Mathcad toolbox has about any math tool you could possibly want, including lots of special functions, vectors and matrices, complex numbers, equation solving, manipulations of inequalities. Further, routines for nonlinear curve fitting via the Levenberg-Marquardt method, cubic splines, fast Fourier and wavelet transforms, Heaviside step functions, etc., are all implemented. In addition, there is a wealth of statistical functions for population statistics, probability distribution, correlation, linear and non-linear regression, automatic conversion of units, and more.

Within Mathcad, 2-D and 3-D graphs are merely plots, and are not particularly polished. In view of the superb graphics capabilities of popular packages such as MATLAB and GAUSS, there is still room for improvement in this area. Changing viewpoints, zooming, and reading out coordinates are controlled via elegant dialog boxes, though, rather than with arcane

commands, and graphics can be imported and exported via the clipboard.

One striking feature of Mathcad PLUS 5.0 is SmartMath, an expert system tool that facilitates combining symbolic and numerical capabilities. Again, the developers of Mathcad thought of convenience. Mathcad's symbolic manipulation features are based on the computer algebra system Maple. Within Mathcad one can perform most symbolic operations for algebra and calculus, compute Laplace, Fourier and Z-transforms and their inverses, but the program has a limited number of Maple tools.

Mathcad comes with interactive materials such as Electronic Handbooks, Application Packs, Function Packs, and Productivity Packs that all aid in customizing Mathcad for different disciplines—including engineering, advanced mathematics, physics and astrophysics, chemistry, statistics, finance and education.

Within Mathcad, an Electronic Handbook can be opened, its contents (consisting of formulas, constants, and diagrams) can be manipulated, cut and pasted into a Mathcad worksheet. Each Application Packs contains an average of 20 complete Mathcad templates for frequently performed calculations.

The Function Packs for digital signal processing, differential equations, data analysis, graphics, and algorithms for numerical analysis, extend the functionality of Mathcad in these areas.

Productivity Packs cover disciplines such as aeronautic, civil, chemical, mechanical and electrical engineering, economics, finance, statistics, mathematics, natural and physical sciences, and education. The MathSoft Electronic Book Sampler that comes with Mathcad 5.0 PLUS, provides samples of these electronic books, which can be bought separately.

To complement Mathcad, MathSoft provides an array of innovative products. Sold separately, the technical professional suite offers enhanced software, together with an Electronic Book that takes you on a guided tour through the mathematical features of Mathcad. A second part offers a comprehensive discussion of the additional features in the PLUS version of Mathcad. Third, a collection of entertaining examples show you in detail the capabilities of Mathcad. Mathcad users can also order electronic versions of many standard reference books, and subscribe to the Applied Mathcad Electronic Magazine, which is published quarterly.

Mathcad emphasizes convenience and attractive output. It is a good tool for doing quick calculations and work-ups, but is less practical for sophisticated work with lengthy input or output. In contrast to MATLAB and GAUSS, it offers only limited programmability and that somewhat restricts what you can do with the program. I would not use it for advanced research projects. If your every day needs are numerical or symbolic, and you care about a snappy user interface, Mathcad is a good choice.

## 6. MATLAB

As soon as you start flipping through the multi-volume set of well-organized documentation that comes with MATLAB, you know you are dealing with an impressive program. After quite

a few additions and enhancements, today's MATLAB is a product of quality and strength. The people at The MathWorks keep up with the swift improvements in computer technology. MATLAB, with its toolboxes, employs fast state-of-the-art algorithms and efficient numerical methods that, with sophisticated management of virtual memory and space barriers, fully exploits the hardware of the nineties.

Recently, when a bug in the Pentium chip was detected, Cleve Moler, chairman and co-founder of The MathWorks, was quick to provide a few lines of MATLAB pseudo code that work around the flaw in the double-precision floating-point arithmetic on the Pentium.

Together with the routines from the IMSL and the NAG numerical libraries, MATLAB is a favorite tool in the scientific computing community, and in the technical and financial industries. MATLAB is best described as a high-powered mathematical engine, fueled by a fully fledged matrix programming language, with as additives extensive integrated mathematical libraries, application development tools, and truly impressive graphics.

With over 500 functions, MATLAB covers all the important linear algebra routines for matrix manipulations, including matrix exponentials and logarithms, SVD and Schur decompositions, LU, QR, and Cholesky factorizations, eigenvalues and eigenvectors, norms, condition numbers, etc. In addition, various special routines are available for handling and visualizing sparse matrices, for convolution, filtering, discrete and inverse FFTs in 1 and 2D. The program has Runge-Kutta algorithms for solving first-order ODEs, interpolation routines based on splines, statistical functions, and much more.

A key feature of MATLAB is its stunning object-oriented graphics: 3-D color surfaces, 3-D mesh and contour plots, 3-D volumetric slice plots, combination surface and contour plots, scatter plots, and 3-D data trajectories. With MATLAB you can create and output publication quality graphics in 2 and 3 dimensions. Graphs are easy. Multiple graphs automatically get a different type of line so they are easily distinguishable. MATLAB's graphics are completely integrated and run in separate windows. The user can interactively visualize and update equations and data in one window, while watching the graphical output in another.

As with O-MATRIX and GAUSS, the well-designed language of MATLAB is interpretive. Consequently, the edit, compile, link and run cycle, familiar to diehard Fortran and C users, is shortened to edit and run. You can specify commands interactively from a terminal, or, include them from an external file, or do both. Multiple command-line statements and function definitions can be written in m-files. The prepackaged m-files of MATLAB provide good examples of how to write custom scripts. I believe that MATLAB's adherence to an open-system philosophy, helped make it into one of the most widely used software packages in the scientific community. MATLAB, just as GAUSS, offers the ability to import code in C and Fortran. With improved file I/O, data can be input to and output from MATLAB in either ASCII or binary format. MATLAB has excellent debugging features and printer support.

The program can be combined with specialized add-on software such as SIMULINK, a powerful interactive workbench that enables you to model, analyze, and simulate nonlinear

dynamical systems. With SIMULINK's Real-Time Workshop and the Accelerator you can automatically generate C code for real-time applications. The MathWorks produces a suite of notebooks, and 15 application-specific toolboxes for signal and image processing, splines, symbolic math (based on Maple), neural networks, statistics, optimization, various control systems,  $\mu$ -analysis and synthesis, Hi-Spec, quantitative feedback theory, systems identification, and chemometrics. The MetaMatrix Toolbox from Atlantic Aerospace, provides a facility for working with data sets with an unlimited number of dimensions within MATLAB. All of these toolkits come with excellent documentation. Since I do not have MicroSoft Word v. 6.0 (or later) for Windows, I could not test the MATLAB notebook suite for Windows and its utilities.

With so much functionality in MATLAB, the reader may be concerned about the learning curve. If you spend a couple of hours with MATLAB, and run some the demonstration files, you'll have most of the information you need to get started on your own problems. Internet users can join a MATLAB news group, subscribe to the quarterly MathWorks newsletter, the electronic 'MATLAB Digest', or access The MathWorks Mosaic home page on the World Wide Web (<http://www.mathworks.com>).

There are a few problems with MATLAB. The interface between MATLAB and its symbolic math toolbox, based on Maple, creates some awkward situations. For example, the syntax needed to display several pictures simultaneously is not obvious. The integration of Maple with MATLAB is not quite seamless. The screen technology and the editor could be improved. For instance, I could not figure out how to get rid of pages of unwanted numbers on the screen, produced unintentionally by a missing semi-colon at the end of a command. This is possibly corrected in the notebook implementation under MicroSoft Word. It is somewhat surprising that MATLAB, which excels in matrix manipulations, fails to provide better matrix editing facilities.

All in all, MATLAB is an excellent software package, indispensable for research and system development within the numerical analysis and engineering communities and the financial industry. If high performance in numerical computation and data analysis is desired, and your application requires modeling, simulation, and the highest quality graphics you should consider MATLAB.

## 7. GAUSS

I was not familiar with GAUSS before the review, but I was quickly struck by the similarities between GAUSS and MATLAB. I can not help but make comparisons of these two fine packages. Their comprehensive programming languages are close, both have large libraries of well-tested linear algebra routines, their visualization, data analysis and graphical capabilities are similar, both have great graphical user interfaces, and they have many numericals features in common. Just like every writer has a different style, the designers of MATLAB and GAUSS use slightly different conventions and notations, nothing to agonize over.

GAUSS is a popular number cruncher and interactive analysis environment, with super strength in the areas of econometrics, biometrics, optimization, and statistics. It has a broader coverage of the distribution functions than MATLAB. Indeed, GAUSS offers the normal, Chi-square, gamma, beta, F and t distributions. On the other hand, GAUSS has no symbolic capabilities, and lacks a few functions such as matrix exponentials and logarithms. Also, if speed is a factor, then GAUSS is hard to beat.

GAUSS includes a nice built-in screen editor with a full range of editing, searching and cut-and-paste commands, whereas MATLAB has only limited screen-editing capabilities. The GAUSS manual, in two volumes, is well-organized, and provides good cross-referencing and an abundance of examples. On line help on functions and operators is also available.

Notwithstanding the fact that it runs under DOS without Windows, what GAUSS can do is impressive. At the core of GAUSS is an extensive library with 400 mathematical functions and procedures for numerical analysis, LINPACK and EISPACK routines for matrix manipulations and sorting, tools for file handling, and more. Due to a complete built-in programming environment, GAUSS's library can be further extended or customized with user-written functions and subroutines. Programmers and analysts can read data from files produced by a text editor, other software or other applications.

Its presentation-quality graphs, produced in 2-D and 3-D are among the best of any mathematical toolboxes, including Mathematica, Maple and MATLAB. While GAUSS allows in-place updating of graphics, it does not have the full range of animation features that MATLAB offers. GAUSS allows one to convert high-resolution graphics into several formats that MATLAB does not support, e.g. binary Lotus PIC files. Unfortunately, neither MATLAB nor GAUSS supports binary formats such as GIF.

Comparing GAUSS to MATLAB, it is hard to say which one is the speediest, most flexible, accurate, powerful or versatile. An in-depth study with benchmarks would be needed for that. MATLAB has been designed for a more engineering/time-series oriented user, whereas GAUSS is more geared toward statistics and econometrics. Additionally, GAUSS offers a wider array of add-on software for estimating a variety of linear and non-linear econometrics models. It is an indispensable tool for financial analysis and forecasting.

Aptech Systems provides a dozen of GAUSS modules for specialized applications, requiring linear regression, linear programming, descriptive statistics, optimization, estimation, times series analysis, curve fitting and nonlinear equation solvers. In addition, an entire line of third party software is available for use with GAUSS-386i, such as spreadsheets, data handling products, numerical algorithms from the LAPACK project, and various packages for design of control systems, non-linear simulation, signal processing, display, and macro-economic modeling.

Summing up, GAUSS is an example of true craftsmanship. For fast numerical computations, modeling, estimation, or sophisticated statistical analysis, GAUSS is a tool worth having around. It is unfortunate that a Windows version of GAUSS is not yet available. GAUSSX, the menu-driven interface designed for econometric analysis, is a step in that direction. With the eventual arrival of Windows 95, and its 32-bit operating system, this all

may change soon.

## 8. Conclusion

Although one can do fairly sophisticated quantitative work with O-MATRIX, Speakeasy, and Mathcad, I consider these packages more suitable for instructional and educational use. This is particularly true for the WYSIWYG product Mathcad from MathSoft, which delivers 30 electronic adaptations of many standard reference works and expository texts, in addition to 10 of its own titles. Time-series commands and statistical functions in Speakeasy make it a good tool for financial analysts and econometricians.

Mathcad PLUS 5.0, with its 32-bit architecture, MATLAB and GAUSS certainly meet the requirements of production and research environments, for they allow you to communicate better with external programs and systems for visualization of very large data sets. These products address the broader range of specific needs of researchers, in areas such as digital image processing, control system design, oil exploration, biomedical applications, to name a few. Users should be forewarned that may have to buy canned add-on software to perform sophisticated and specialized tasks. Although equipped with symbolic capabilities, neither MATLAB, nor Matcad can replace major symbolic packages such as Mathematica, Macsyma and Maple.

If you are looking for simple plots of rather small data sets, or if you want to illustrate concepts, O-MATRIX, Speakeasy and Mathcad are for you. If you deal with extensive data sets, need high degrees of accuracy or speed, powerful data analysis in a strong programming environment, or quality graphs for publication, MATLAB and GAUSS are products to consider.

Also, MathPlus is a nice addition to the list of symbolic packages for the PC. It offers an elegant user interface, yet it lacks the symbol crunching capabilities a researcher might desire.

Willy Hereman  
Associate Professor  
Department of Mathematical and Computer Sciences  
Colorado School of Mines  
Golden, CO 80401-1887, USA