KU ant Guides

A Template for KUant Guides written with LaTeX

Johnson, Paul E.
<pauljohn@ku.edu>

http://quant.ku.edu

Guide No.

XXXX

This is abstract of the template LaTeX document for the preparation of KUant Guides (http://quant.ku.edu). The abstract should include a general explanation of the guide's topic.

This template is prepared as a LyX document, KUant-template.lyx, but it is not necessary to use LyX if one wants to create a LaTeX document. One can export a LaTeX document from LyX, and so LaTeX users who prefer to write in "pure, clean, authentic, and more-virtuous" LaTeX can do so. (Look for the file KUant-template.tex.)

The design of the "top matter," the part above this abstract, should not be changed by the author. The author should only see fit to change

1. The title of the KUant Guide

2. The author's name & email

3. The number of the KUant Guide

This is the end of the abstract. The part after this is supposed to be presented in two column format.

A KUant Guide is a brief (less than 10 pages) illustration of software usage for a particular task.

The most difficult part of using LaTeX to prepare KUant Guides is that the style guideline requires a two column document. The two column document is implemented here with the LaTeX package multicol.

The multicol package is easy to use, but the two column design worsens a couple of problems that can arise in all LaTeX documents. Simply put, if the user writes a long equation (or inserts a large graphic), the equation (or graphic) is not automatically resized to fit in the column. A figure or an equation that is too large for the space will "flow" out of the column borders. That's a common LaTeX problem, but experienced users are aware of it and, when the document has only one column, it is easier to navigate.

A related problem is that floating figures, a staple of ordinary LaTeX document preparation, are not workable. One can insert a floating figure in order to obtain a figure label and number, but the position must not be allowed to float freely. The user should use the "place here definitely" rubric to force the figure into place.

It is not necessary to use floats for graphic images. A graphic image can be inserted directly, but it is important for the author to make sure that the graphic does fit within the space allowed by the column. That can be done either by editing the graphic file so that its size is small–say 3 inches or less–or by using the graphic sizing options in the insert graphic option. A figure that is inserted into a floating figure may not cause trouble. It should be automatically relocated by LaTeX.

Equations

The issue with equations is maddening for all LaTeX novices. After preparing a beautiful, elaborate formula, the user is dismayed to find that, in the pdf output, the equation overflows into the margins and off the page. It is just a (frustrating) fact of life. LaTeX does not enforce line breaks within equations, so equations can run off the edge of the page, or, in this case outside the edges of the column.

Short equations are going to come out fine, as in

$$G[x] = \frac{1}{2\sqrt{N}} e^2 dx. \tag{1}$$

Authors who want longer equations must manually introduce line breaks.

$$F(p_i) \;=\; \int_{-\infty}^{\infty} G(x)dx +$$

$$\sum_{i=1}^{N} f(x_i)\Gamma(k(m+1)) +$$

$$\begin{pmatrix} 17 \\ 9 \end{pmatrix} p^N (1-p)^{N-k} \qquad (2)$$

This is one of the few elements of LaTeX that really goes against the general LaTeX philosophy of leaving the styling and formatting of output to the LaTeX engine and the publisher's preferred class files.

What goes wrong if the user writes all of that on one line? The output flows into the margin and over the edge.

$$F(p_i) = \int_{-\infty}^{\infty} G(x)dx + \sum_{i=1}^{N} f(x_i)\Gamma(k(m+1)) + \begin{pmatrix} 17 \\ 9 \end{pmatrix} p^N (1-p)^{N-k}$$

$$(3)$$

As far as I understand LaTeX, this is an inherent problem (feature? bug?) with displayed equations. LaTeX lets them flow off the edge of the paper. Authors must re-format long equations to fit within the space allowed.

Graphics

Now, suppose we want to insert a graphic? Ordinarily, in LaTeX documents, we would not insert graphics at a particular spot. Instead, we would wrap the graphic inside floating figure environments.

This is discussed in the LyX User Documentation. Graphics can be "stuck" right here. If we don't use figure floats, what happens? Suppose we have a 6 inch wide image of a jayhawk. We don't have to manually shrink the image, we can re-scale it on the fly. LyX has a pull down Insert -> Graphics menus for that. In LaTeX, the markup would be something like:

```
\includegraphics[width=3in]{jayhawk}
```

We write "jayhawk" and the LaTeX engine will choose among files with names like jayhawk.png, jayhawk.eps or jayhawk.png. It does not wrap it in a "figure" thing, it has no caption. There's no Figure title, no Figure number. Its just a picture, that is pasted in right here.

All good, as long as the image is small, say 3 inches wide, or less, it fits inside the column.

LaTeX won't stop us from inserting a graphic that does not fit. If we insert a 4.5 inch wide image, it will flow off the edge, just like the long equation did.

Thus authors are behooved to make sure that graphics inserted are SMALL (note my irony of making the word "small" large).

Why is that not good enough? Well, as explained in the documentation for LyX, inserting graphics without wrapping them inside "figure floats" is considered somewhat gauche. We'd rather have a numbered Figure, something we can refer to without saying "in the picture that was inserted two paragraphs ago..."

Unfortunately, ordinary figure or table floats will not work inside multicolumn documents. The content of the figures just disappears from the output. It is as if the LaTeX processing system is saying to us, "that was such an amateurish thing to do, we are going to treat it as a joke and ignore it."

But all is not lost. Ordinary floating figures don't work correctly inside columns, but we have *two reasonable alternatives.*

Option 1. Insert a small figure float and tell LaTeX to place it "here". In the LaTeX markup, that's the H placement option. Some people say H means "right here." Some say "here definitely".

In LyX, one can use the Insert -> Float -> Figure menu to create the float. Then one inserts the graphic inside that object. Then right click on the floating figure and choose the "here definitely" placement option.

In raw LaTeX, the markup that corresponds would be

```
1  \begin{figure}[H]
2  \includegraphics[width=3in]{jayhawk}
3  \caption{"Here Definitely" Floating a
       Jayhawk\label{jayfig}}
4  \end{figure}
```

Note the capital H. That's the "right here" "definitely" "I really mean it, I'm not joking" part.

Figure 1: "Here Definitely" Floating Jayhawk

The magic H can work for figure floats that contain any kind of content, as long as it is narrow enough to fit. LyX has a dialogue to insert various kinds of documents from various programs as "external material." Here I demonstrate an XFig diagram and, because it

is inserted in a figure environment. This is in Figure 2. The automatic numbering is very convenient.
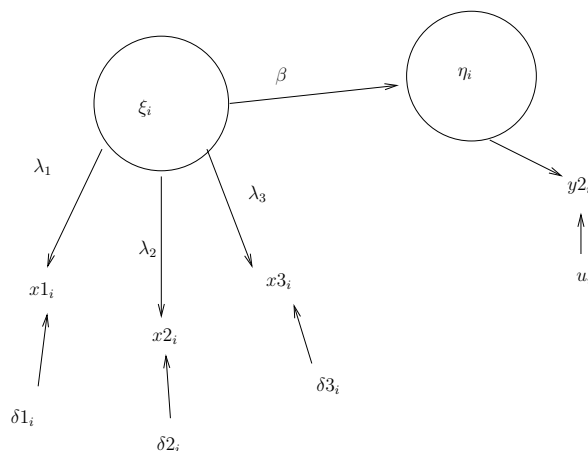

Figure 2: An Xfig Drawing from External Material Dialogue

If we were writing this in raw LaTeX, the markup for that figure would be

```
1  \begin{figure}[H]
2  \resizebox{3in}{!}{\input{concept1.pdftex_t}}
3  \caption{An Xfig Drawing from External
       Material Dialogue}
4  \end{figure}
```

Option 2. Use a special figure environment. This will allow inclusion of large figures, but they will float down to appear on a separate page, not in the 2 column format.

There is no LyX menu for this special figure, it must be inserted as raw LaTeX. The only difference is that the environment is now `figure*`, rather than `figure`.

Since that carries the label "xfigstar", we can refer to it and obtain automatic numbering. For example, if we insert the crossreference as \ref{xfigstar}, we can urge the reader to look at Figure 3.

Code Listings

Finally some good news! There is a LaTeX package called Listings and it is extraordinarily customizable. Document-wide preferences can be set in the document preamble, as I have done in this document. Individual code listings can be adjusted, of course.

Computer code examples should be inserted within program listings environments. Those environments WILL wrap long lines.
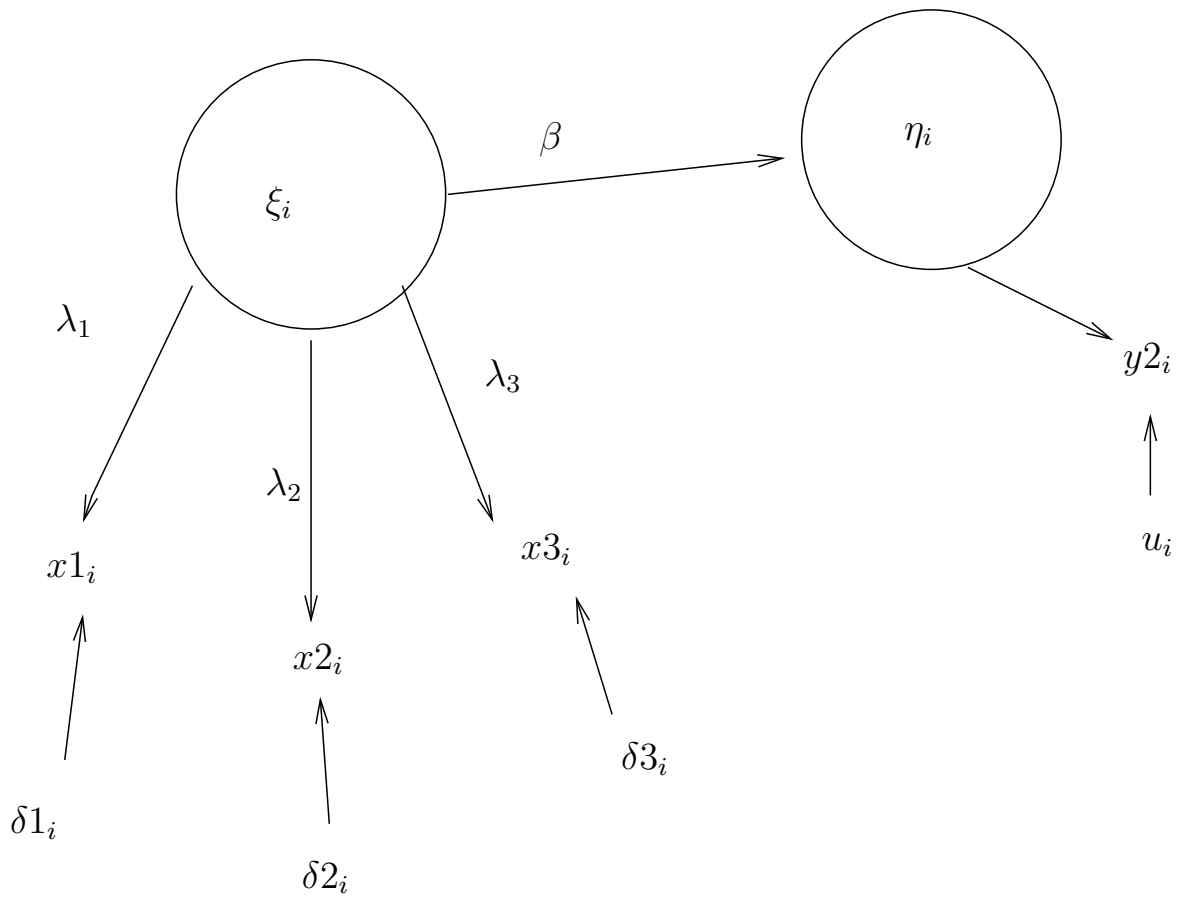
Figure 3: A Figure* environment floats to another page

```
1  x <- rnorm(100, m = 343, s = 232)
2  y <- rgamma(100, lamda = 534)
3  plot(x, y, xlab = "This is the x variable", ylab = "This
       is the y variable", xlim = c(0, 100), y = c(-100,
       2000)
```

I fiddled it quite a bit, seems ok. The listings are almost infinitely customizable. One can set the language, which draws in a pre-set collection of choices, and then the style can be further customized (for, say, R code input and R text output). There are predefined language settings for R, SAS, and many other languages.

Conclusion

By far, the most difficult part of this exercise is the two-column design, which complicates the inclusion of equations and figures. But, once one becomes aware of the issue (it was a complete surprise to me), the problem can be managed.

I wanted to insert some bibliographic citations, so that future users might have a fully working example. So, without further ado (Aitkin, 1999), and Albert (2007, p. 57), as well as others (Jackman, 2009; McCullagh, 2008; McCullagh & Nelder, 1983).

The other problem I need to work on is the usage of Sweave for R document preparation. I'm not sure how well that will integrate into this format.

References

Aitkin, M. (1999). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, 55(1), 117–128.

Albert, J., Ed. (2007). *Bayesian Computation with R*. New York, NY: Springer New York.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. Chichester, UK: John Wiley and Sons.

McCullagh, P. (2008). Sampling bias and logistic models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 70(4), 643–677.

McCullagh, P. & Nelder, J. A. (1983). *Generalized Linear Models*. Number 37 in Monographs on statistics and applied probability. London: Chapman and Hall.