Fonts in Plots

Paul Murrell

July 5 2006
Selecting Fonts in R

Font size
The size of text can be controlled via absolute point size and through the use of a character expansion multiplier.

text(x, y, "whatever", cex=2)
grid.text("whatever", gp=gpar(cex=2, fontsize=10))

Font face
The style of text can be chosen from plain (or roman), bold, italic, bold-italic, or “symbol”.

text(x, y, "whatever", font=2)
grid.text("whatever", gp=gpar(fontface="bold"))
Font family

The overall font can be chosen from sans-serif, serif, or monospaced.

text(x, y, "whatever", family="serif")
grid.text("whatever", gp=gpar(fontfamily="serif"))

Custom font family

A new font can be defined using a font database which maps a family name to font details.

pdfFonts(cmss=Type1Font("Computer Modern",
c("cmopsfont/afm/cmss10.afm",
"cmopsfont/afm/cmssbx10.afm",
"cmopsfont/afm/cmsssi10.afm",
"cmopsfont/afm/cmssbx10.afm")))
Why Worry?

**Legibility**

The choice of font is important for legibility: a sans serif font is more likely to be legible in small pieces of text, especially on screen.

![Legibility is Good with a Sans Serif Font](image1)

![Legibility is Not So Good with a Gothic Font](image2)
Consistency

The choice of font can be important for consistency: it can look nicer and more professional if the font used in figures matches the font used in the main text of a book or report.

This is Helvetica

This is Computer Modern Sans Serif
Why Worry?

Special characters

The choice of font can be forced upon you: sometimes you need a character that is only available in a particular font.
Localisation

The choice of font can be cultural: some languages require a particular font because they use a completely different set of characters.

Hello World

你好 世界
Aesthetics

The choice of font can be a matter of taste: sometimes you need the right font to make text look nice.

At first sight it must seem intolerably degrading for Zen — however the reader may understand this word — to be associated with anything so mundane as archery. Even if he were willing to make the big concession, and to find archery distinguished as an “art,” he would scarcely feel inclined to look behind this art for anything more than a decidedly sporting form of prowess. He therefore expects to be told something about the amazing feats of Japanese trick-artists who have the advantage of being able to rely on a time-honored and unbroken tradition in the use of bow and arrow. For in the Far East it is only a few generations since the old means of combat were replaced by modern methods, and familiarity in the handling of them by no means fell into disuse, but went on propagating itself, and has since been cultivated in ever widening circles. Might one not expect, therefore, a description of the special ways in which archery is pursued today as a national sport in Japan?
Mathematical Formulas in Plots

The champion

\( \text{TEX} \) is the gold standard for formatting mathematical formulas.

The pretender

\( R \) provides a mechanism for formatting mathematical formulas, which attempts to emulate \( \text{TEX} \). How does it do?

\[ \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

\[ \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]
A token effort

We can do a little bit better with a serif font and some italics.

\[
\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]
The Problem

Confession

- A minor problem is that we are using a Times font not \( \text{T}_{\text{E}}\text{X}'s \) Computer Modern Math Italic font. This can easily be fixed by specifying the appropriate font.

- The main problem is that we are using the Adobe Symbol font not \( \text{T}_{\text{E}}\text{X}'s \) Computer Modern Math Symbol font. This is not so easy to fix because just using the appropriate font fails miserably.

\[
\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

\[
\int \sqrt{2} e^{-\frac{x^2}{2f^2}}
\]
The Problem

The gory details

- The underlying problem is that R assumes that the symbol font uses the Adobe Symbol Encoding, but the Computer Modern fonts use special \TeX-specific encodings.
- Another problem is that the Computer Modern Math Symbol font does not contain (anywhere near) all of the symbols in the Adobe Symbol Encoding.

| \ suchthat | \ similarequal |
| ( parenleft | \arrowdblleft |
| ) parenright | \arrowdblright |
| * asteriskmath | \arrowdbup |
| + plus      | \arrowbldown |
| , comma     | \arrowdblboth |
| - minus     | \arrownortwest |
The Solution

What would any self-respecting Statistician do?
Create a new font of course!

How do you create a new font?

1. Pay lots of money to someone with lots of talent.
2. Use a fancy program and spend many hours fiddling.
3. Cheat.

Source: FontForge editing tutorial
http://fontforge.sourceforge.net/editexample.html
The Solution

The structure of a Type 1 font

- A Type 1 font consists of a font metric file, `fontname.afm`, which is a text file, and a font outline file, `fontname.pfb`, which is a binary file.
- The `t1disasm` program (from `t1utils`) turns `fontname.pfb` into `fontname.raw`, which is a text file.

```
C 39 ; WX 666.667 ; N suchthat ; B 83 -40 583 540 ;
C 40 ; WX 388.889 ; N parenleft ; B 99 -250 331 750 ;
C 41 ; WX 388.889 ; N parenright ; B 57 -250 289 750 ;
C 42 ; WX 500 ; N asteriskmath ; B 65 34 434 465 ;
C 43 ; WX 777.778 ; N plus ; B 56 -83 721 583 ;
C 44 ; WX 277.778 ; N comma ; B 86 -193 203 106 ;
C 45 ; WX 777.778 ; N minus ; B 83 230 694 270 ;
...
```

```
... /minus {
83 7000 9 div
hsbw
230 40 hstem
0 611 vstem
578 230 rmoveto
14 19 0 20 hvcurveto
20 -19 0 -14 vhcurveto
-545 hlineto
-14 -19 0 -20 hvcurveto
-20 19 0 14 vhcurveto
closepath
endchar
} ND
...
```
A new font can be created from existing fonts by simply reordering and combining character metric and outline information (in text formats).

The full Adobe Symbol Encoding requires characters from:

- Computer Modern Roman
  (e.g., space, left and right parentheses)
- Computer Modern Math Symbol
  (e.g., core math symbols, minus)
- Computer Modern Math Italic
  (e.g., less than, greater than, Greek symbols)
- Computer Modern Math Extension
  (e.g., arrows, math operators)
- AMS Math Symbol [A] (therefore, angle, lozenge)

The `t1asm` program (from `t1utils`) turns `fontname.raw` into `fontname.pfb`. 
The Solution

Computer Modern Adobe-Symbol-Encoding Math Symbol

- The font can be obtained from http://www.stat.auckland.ac.nz/~paul/R/CM/CMR.html as cmsyase.afm and cmsyase.pfb.

Limitations

- The proposed solution involves a Type 1 font, which covers PDF and PostScript output.
  - A Windows “port” is possible in theory, but has not worked yet in practice.
- The new font is still missing some characters and some could do with a bit more polish.
Almost perfect


\[
\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]
Another example

Conclusions

Summary

• Fonts are important in plots.
  • Legibility
  • Consistency
  • Special characters
  • Localization

• When producing mathematical formulas in plots, it is now possible to use the fonts that \TeX uses.
  • Even beauty matters
Acknowledgements

Fonts used in this talk

- The Hershey vector font Gothic English Triplex (originally incorporated from the GNU plotutils package; now distributed as part of R).
- The Computer Modern PostScript Fonts (Adobe Type 1 format)
- The AMS PostScript Fonts (Adobe Type 1 format)
- The Type 1 CM-based fonts for Latin, Greek and Cyrillic, with Latin 1 encodings, from the cm-lgc TeX font package.

Other stuff used in this talk

- The typesetting example used the first paragraph of “Zen in the Art of Archery” by Herrigel (taken from Jim Heffron’s TUGboat article “Why TeX?”).