**note**

**@46**

** ▾**

**6**views

[stop following](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

the problem with mmplot: email me your code

Could someone who was experiencing the error with mmplot email me their code (for fitting the model as well as the mmplot itself).

I’m having a hard time reproducing the error from what I remember from office hours from this morning.

thanks

-BJ

[hw7](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

[**edit**](https://piazza.com/class/kt1x3afkm0yfp?cid=46)·[good note](https://piazza.com/class/kt1x3afkm0yfp?cid=46)0

Updated 3 hours ago by

Brian Junker

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



**[Emily Zeng](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [2 hours ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

i think the issue has been solved! basically, when creating the model with glm, it should look like glm(y ~ x + x2 + etc, fam = binomial, data = data) instead of glm(y ~., data = data).

[helpful!](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Brian Junker](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [2 hours ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

Thanks for the update!

Yes, you do have to provide the family for glm’s.

If anyone else has questions about this, let me know.

best,

-BJ

[good comment](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0

Reply to this followup discussion

[Actions](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

Resolved Unresolved

**@46\_f2**



**[Sifeng Li](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [2 hours ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46) Hi I guess I still have problem after I added the family for glm’s for model in part c, and the error is still out of vertex space!

[helpful!](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Brian Junker](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [2 hours ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

OK, can you paste your code (reading the data, fitting the model, and running mmplot) into a reply here so I can take a look at it?

thanks

-BJ

[good comment](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Zixuan Jin](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [1 hour ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

nes <- read.dta("nes5200\_processed\_voters\_realideo.dta")

nes\_1976 <- nes %>%
  filter(year == 1976)
nes\_1976$rep\_pres\_intent <- as.factor(nes\_1976$rep\_pres\_intent)
myvars <- c("ideo7", "black", "female", "rep\_pres\_intent", "presapprov", "age",
            "educ1", "urban", "income", "union", "perfin1")

newdata <- nes\_1976[myvars]

data.num <- newdata %>%
  mutate(ideo7=as.numeric(ideo7)) %>%
  mutate(black=as.numeric(black)) %>%
  mutate(female=as.numeric(female)) %>%
  mutate(presapprov=as.numeric(presapprov)) %>%
  mutate(educ1=as.numeric(educ1)) %>%
  mutate(urban=as.numeric(urban)) %>%
  mutate(income=as.numeric(income)) %>%
  mutate(union=as.numeric(union)) %>%
  mutate(perfin1=as.numeric(perfin1))

lm2 <- glm(rep\_pres\_intent ~., data = data.num, family = binomial)
summary(lm2)

mmplot(lm2)

[helpful!](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Brian Junker](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [56 minutes ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

Thanks. I can reproduce the error; I will try to get a fix or workaround out soon.

-BJ

[good comment](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Zixuan Jin](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [55 minutes ago](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

Thank you so much!

[helpful!](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0



**[Brian Junker](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [Just now](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

Short answer:

Try

mmplot(lm2,locfit.control=list(nn=1))

Long answer:

mmplot uses the function locfit from library(locfit) to fit smooth nonparametric regression functions to (a) the raw data and (b) the predictions from the glm model.

help(mmplot)

led me to the parameter locfit.control through which you can pass a list of parameters to the underlying locfit.raw function (that actually does the calculations), to adjust the smoothness of the fit. Generally speaking, locfit.raw requires more memory for less smooth fits (which is what I would expect for a local polynomial spline model, for example), and less memory for more smooth fits. The thing that caused the error was that the default “smoothness” for locfit required too much memory for some of the variables in the marginal model plot.

help(locfit.raw) gave me some parameters I could tweak to adjust the smoothing, but it also led me to the function lp(), and help(lp) led me to some more. Some key parameters that affect smoothing and memory use are:

maxk -- seems to reserve maximum memory available for the fit

 h -- a "bandwidth" parameter: larger h --> smoother fit

 nn -- a "nearest neighbor" parameter: larger nn --> smoother fit

* The default value for maxk is 100, but I could not fix the memory problem by increasing maxk, even to 1000.
* The default value for h is 0, but I could not get the right tradeoff between blowing out the memory and having too smooth a fit by increasing h (note: if you set h='aic', then locfit.raw will choose h by minimizing AIC, but the h you get is around 14, which makes nonparametric regressions that just look too smooth to my eye).
* The default value for nn is 0.7 (whatever that means!) and by trial and error with some values larger than 1 (both a little larger and a lot larger) and some values between 0 and 1, I found what seems like a reasonable tradeoff between good memory usage and not too much smoothness at nn=1.

Finally, for this part of the problem it’s not strictly necessary to use the mmplot function, but the results are interesting , and so it turned out to be a good idea to try (despite the difficulty of figuring out how to tweak the function).

hope this helps,

-BJ

[good comment](https://piazza.com/class/kt1x3afkm0yfp?cid=46) 0

[Actions](https://piazza.com/class/kt1x3afkm0yfp?cid=46)



**[Brian Junker](https://piazza.com/class/kt1x3afkm0yfp?cid=46)**

 [Just now](https://piazza.com/class/kt1x3afkm0yfp?cid=46)

PS - I also ended up trying to find some papers that explain the theory behind locfit. It turns out that there is a book,

Loader, C. (2006). *Local regression and likelihood*. Springer Science & Business Media.

as well as several papers and talks.

Dr. Catherine Loader has done some interesting and valuable work in nonparametric regression and density estimation, but also has an interesting life story: she started out as Clive Loader.

all best,

-BJ