Meeting with Dom 20210322, and subsequent work

Lena Morrill

April 6, 2021

Contents

1 Convergence and optimisers

Fixed effects DM   When removing the coefficient of 100- for the DM likelihood, I get convergence. Otherwise the gradient goes to 0 and I get NA for every single standard error of the parameters.

Diag RE effects DM   kidney renal clear cell RCC now converges (it didn't use to) when removing the 1000 coefficient, and using nlminb.

Full RE effects DM   Removing the coefficient of 1000, still no convergence for kidney renal clear cell RCC, even if increasing number of iterations, and using nlminb, but I don't get the message about gradient too steep.

When using a high value for the initial log lambda (so that it's closer to a multinomial), it doesn't converge either.

When initialising the covariances around 0, and the u at 0, I get that the Hessian of fixed effects was not positive definite., but no NAs.

Not including covariances for signatures with many zero exposures   That gives good convergences!

2 Bias and coverage

The beta intercept estimates are biased (to high values) in detriment of the standard deviation of the random effects. The overdispersion parameters are also biased to higher values.

1. Using GenerationCnorm, set 2. **Note! The random effects where actually completely correlated in this generation, but the covariances to assess bias were said to be zero, so the results are incorrect!**

   (a) diagREDM: high beta intercept, good beta slope, low lambda, low standard deviation RE

   (b) fullREM: high beta intercept, good beta slope, high standard deviation RE, high covariances RE

   (c) fullREDM single lambda: high beta intercept, good beta slope, good covariances RE, high lambda, high standard deviation RE

   (d) diagREDM single lambda: high beta intercept, good beta slope, high lambda, low standard deviation RE

   (e) diagREM: high beta intercept, good beta slope, high standard deviation RE

   (f) fullREDM: high beta intercept, good beta slope, good covariances RE, high lambda, low standard deviation RE

2. Using GenerationCnorm, set 1

   (a) fullREDM: high beta intercept, good beta slope, good covariances RE, high lambda, low standard deviation RE

3. Using GenerationDMFE1, set 1 and 2, and varying lambda (overdispersion): even in the case where there are no random effects there is a bias in the beta intercept (which is over-estimated). Lambda is too over-estimated. In the second beta set (set 2) not even all the slope betas are unbiased. Set 3 is as bad as set 2

   - With fewer categories ($d = 3$):

Dataset	Set idx	Model (full)	RE structure	Model	Overdispersion structure[1]	Beta intercept	Beta slope	Lambda (overdisp.)	Std RE	Cov RE
GenerationCnorm[2]	1	fullREDM	fullRE	DM	DL	H	G	H	L	G
	2	fullREDM	fullRE	DM	DL	H	G	H	L	G
	2	diagREDM	diagRE	DM	DL	H	G	L	L	-
	2	fullREDM SL	fullRE	DM	SL	H	G	H	H	G
	2	diagREDM SL	diagRE	DM	SL	H	G	H	L	-
	2	fullREM	fullRE	M	-	H	G	-	H	H
	2	diagREM	diagRE	M	-	H	G	-	H	-
GenerationCnormdiagRE	2	fullREDM SL	fullRE	DM	SL	H	G	L	H	H
GenerationMGnorm	1	fullREM								
GenerationMGnorm	1	diagREM								

In singlelambda, the biases are the same in both datasets???

---

[2]DL: Double overdispersion parameter (one for each group). SL: Single overdispersion parameter (shared).

[2]Incorrect model! The intercepts are all the same, so there is a perfect correlation! This is not what I said in the assessment of the bias