SNP agging conditions

Stas

10 January, 2019

Let Z_o be the observed Z scores for the joint 3-SNP model, with $Z_o \sim N(Z_M, \Sigma)$, $Z_M = \Sigma Z_J$. We are interested in evaluating plausibility of situations in which we would be forced to conclude that SNP 3 is the causal one. Two situations are possible: SNPs 1 and 2 are casual with $Z_J = z_J = (\zeta_1, \zeta_2, 0)'$, or SNP 3 is causal with $Z_J = \tilde{z}_J := (0, 0, \tilde{\zeta})'$. Likelihood of SNP 3 being causal is larger than the likelihood of SNPs 1 and 2 being causal when

$$2(\tilde{\zeta} z_3^o - \zeta_1 z_1^o - \zeta_2 z_2^o) + \zeta_1^2 + \zeta_2^2 - \tilde{\zeta}^2 + 2\zeta_1\zeta_2 r_{12} > 0. \tag{1}$$

SNPs 1 and 2 are causal

We assume $\zeta_1 = \zeta_2 := \zeta$ and $\tilde{\zeta} = \zeta(r_1 + r_2)$ and calculate the probability of the condition (1) above being satisfied as

$$1 - \Phi\left(\frac{1}{2}|\zeta|\sqrt{2(1 + r_{12}) - (r_1 + r_2)^2}\right).$$

Additionally conditions

$$\begin{gathered} -2r_1 r_2 r_{12} + r_1^2 + r_2^2 + r_{12}^2 \leq 1 \\ r_1 + r_2 < \sqrt{2(1 + r_{12})}. \end{gathered}$$

must be satisfied.

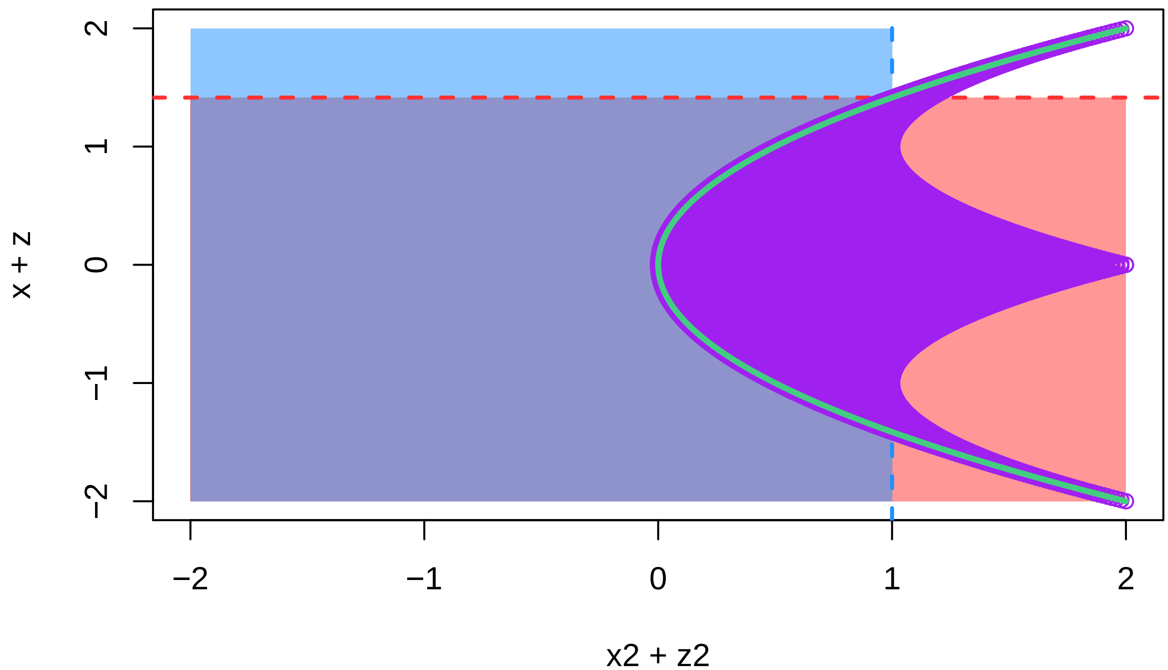
Suppose $r_{12} = 0$. Then the above conditions reduce to

$$\begin{gathered} r_1^2 + r_2^2 \leq 1 \\ r_1 + r_2 \leq \sqrt{2}. \end{gathered}$$

```
r <- seq(-1, 1, by = 0.01)
dat <- expand.grid(r1 = r, r2 = r)

plot(c(-2, 2), c(-2, 2), type = 'n', xlab = 'x2 + z2', ylab = 'x + z')
rect(-2, -2, 2, sqrt(2), col = alpha('firebrick1', 0.5), border = NA)
rect(-2, -2, 1, 2, col = alpha('dodgerblue', 0.5), border = NA)
abline(h = sqrt(2), lty = 2, col = 'firebrick1', lwd = 2)
abline(v = 1, lty = 2, col = 'dodgerblue', lwd = 2)

points(dat$r1^2 + dat$r2^2, dat$r1 + dat$r2, col = "purple")
lines(r^2 + r^2, r + r, col = "seagreen3", type = 'l', lwd = 3)
```



2
1
0
−1
−2
x + z
−2
−1
0
1
2
x2 + z2