AMS 268 – Homework 2
Mickey Warner

1 g-prior

We set $g = \max(n, p^2)$, which became $p^2 = 400$ in both cases. 95% equal-tailed probability intervals for β_1 and β_{10} posteriors are given in Table 1. In both cases ($n = 50$ and $n = 200$), the interval for β_1 does not contain zero and the interval for β_{10} does contain zero. So we would reject the hypothesis that $\beta_1 = 0$ and fail to reject $\beta_{10} = 0$.

n	β_1		β_{10}	
50	2.56	3.40	-0.93	0.13
200	2.76	3.14	-0.33	0.16

Table 1: 95% credible intervals for β_1 and β_{10}.

2 Random Forest and BART

Model	n	p	ntree	Noise?	Coverage	Length	MSPE
RF	200	200	10		0.96	28.03	34.34
RF	200	200	10	Added	0.96	29.41	40.64
RF	200	200	500		0.99	45.43	31.26
RF	200	200	500	Added	0.99	45.60	31.81
RF	500	100	10		0.97	29.86	41.78
RF	500	100	10	Added	0.96	30.20	45.93
RF	500	100	500		0.99	46.62	31.22
RF	500	100	500	Added	0.99	45.97	31.65
BART	200	200	10		0.78	16.15	46.44
BART	200	200	10	Added	0.69	12.50	37.12
BART	200	200	500		1.00	14.39	1.32
BART	200	200	500	Added	1.00	14.74	1.44
BART	500	100	10		0.43	9.87	55.44
BART	500	100	10	Added	0.36	7.82	53.90
BART	500	100	500		1.00	12.34	1.27
BART	500	100	500	Added	1.00	14.02	1.96

Table 2: Results from fitting random forest and BART models.

Table 2 shows the results for each of the 16 data generation and model fitting scenarios. In the table, Coverage is the proportion of prediction intervals that contained the true value, Length is the average length of all prediction intervals, and MSPE is mean-squared prediction error

$$MSPE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2.$$

In all cases, the same data used to fit the models was used to make predictions. Prediction intervals for the random forest are calculated using the individual predictions from each tree.

Notice that the coverage for random forests is typically on the high side, except for when few trees are used. But in these cases the MSPE suffers. Paradoxically, the average interval length was greater when more trees were used, which also accounts for the greater coverage.

BART had better predictive performance (when more trees were used) and smaller intervals on average than random forest. Though, the coverage is suspect. With few trees, we are way below the mark. With more trees, the intervals are too wide. Such intervals are still smaller compared to random forest, while having comparable coverage.

BART appears to be superior to random forest, provided enough trees are used. Prediction intervals for BART are overly conservative, given the universal coverage.