Week6 Solutions to Theoretical Exercises

Kristoffer Olesen

March 2021

1 Exercise 1

Using first tip 2 and then tip 1 we find:

$$\begin{aligned} Var(\frac{1}{B}\sum_{i=1}^{B} B_i) &= \frac{1}{B^2} Var(\sum_{i=1}^{B} B_i) && (1) \\ &= \frac{1}{B^2}\sum_{i=1}^{B} Var(B_i) && (2) \\ &= \frac{\sigma^2}{B} && (3) \end{aligned}$$

As B approach infinity the variance will go to zero.

2 Exercise 2

Using multiple applications of tip 1 and symmetry of the covariance we find:

$$
\begin{aligned}
Var(\frac{1}{B}\sum_{i=1}^{B} B_i) &= \frac{1}{B^2} Var(\sum_{i=1}^{B} B_i) && (4)\\
&= \frac{1}{B^2} Var(B_1 + \sum_{i=2}^{B} B_i) && (5)\\
&= \frac{1}{B^2}\left(Var(B_1) + 2 \cdot CoV(B_1, \sum_{i=2}^{B} B_i) + Var(\sum_{i=2}^{B} B_i)\right) && (6)\\
&= \frac{1}{B^2}\left(Var(B_1) + 2\sum_{i=2}^{B} CoV(B_1, B_i) + Var(\sum_{i=2}^{B} B_i)\right) && (7)
\end{aligned}
$$

We can use the same calculation to withdraw each single $Var(B_i)$ and end up with:

$$
\begin{aligned}
Var(\frac{1}{B}\sum_{i=1}^{B} B_i) &= \frac{1}{B^2}\left(B\sigma^2 + 2\sum_{i=1,j>i}^{i=B-1} CoV(B_i, B_j)\right) && (8)\\
&= \frac{1}{B^2}\left(B\sigma^2 + 2\sum_{i=1,j>i}^{i=B-1} \rho\sigma^2\right) && (9)\\
&= \frac{1}{B^2}\left(B\sigma^2 + 2\rho\sigma^2 \sum_{k=1}^{k=B-1}\right) && (10)\\
&= \frac{1}{B^2}\left(B\sigma^2 + 2\rho\sigma^2 \frac{(B-1)(B-1+1)}{2}\right) && (11)\\
&= \rho\sigma^2 + \frac{\sigma^2(1-\rho)}{B} && (12)
\end{aligned}
$$

We see now as B approach infinity the variance will approach a constant determined by the correlation and variance of each individual tree.

3 Exercise 3

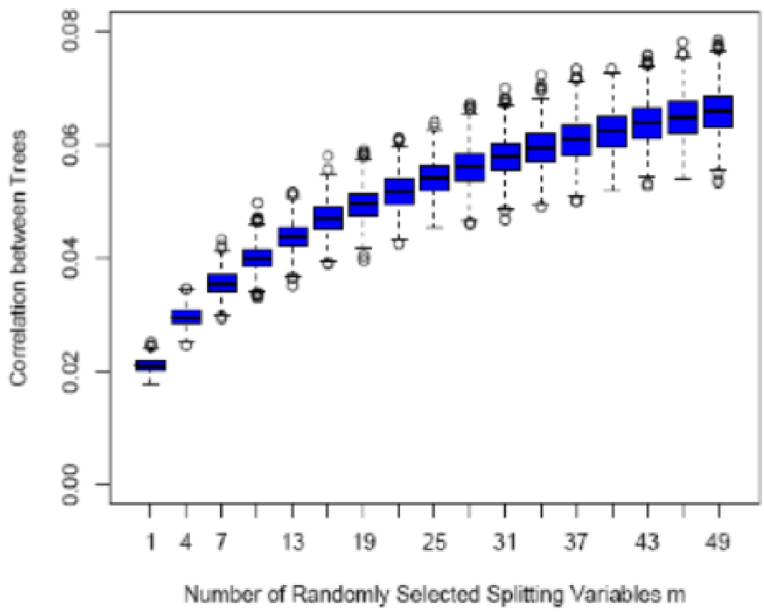

Figure 1: Correlations between pairs of trees drawn by a random-forest regression algorithm, as a function of m. The boxplots represent the correlations at 600 randomly chosen prediction points x.

From the figure we see the correlation between trees decreases as the number of selected splitting variable decreases. From the result in exercise 2, this means that the variance of B trees decreases as the number of selected splitting variable decreases.