*Demonstration of the smallest tip equals to the minimum of the eigenvalues by a simple 3 taxon case*

The tip lengths for species and are , and , respectively. So the shortest tip branch length is (for species and ). The length matrix for the tree can be represented as following

Claim: is the smallest eigenvalues.

Proof: Let be an the eigenvalue for , then we have where is an 3 by 3 identity matrix. Solving the equation step by step, we have

then we have 3 eigenvalues .

Obviously the smallest eigenvalue is which is the shortest tip length.

*General proof: smallest tip is the minimum of the eigenvalues.*

Note that given an ultrametric tree of tips, there exists a unique strictly ultrametric matrix (Nabben and Varga 1994) to represent the species relationships.

Again let be the smallest tip length. Then has at least two identical columns or rows where is an by identity matrix. This implies that

which implies that is an eigenvalue of . The next step is to show that is the smallest eigenvalue for the eigenvalue set of .

*Claim*:

*Proof*: If is an eigenvalue of satisfying , it suffices to show that . Suppose to the contrary that . Consider the matrix , then is still an strictly ultrametric matrix which is always invertible (see Nabben and Varga 1994, and Corollary 6.2.27 in Horn and Johnson 1985). We have which implies . This consequence indicates that can not be an eigenvalue of which contradicts to the assumption of Therefore, is the smallest eigenvalue of .

*Why tiny branch lengths cause ill-condition matrix*

The problem of having ill-conditioned matrix comes from the tiny branch lengths.

We can illustrate this issue using a simple example of two taxa shown in the following

1

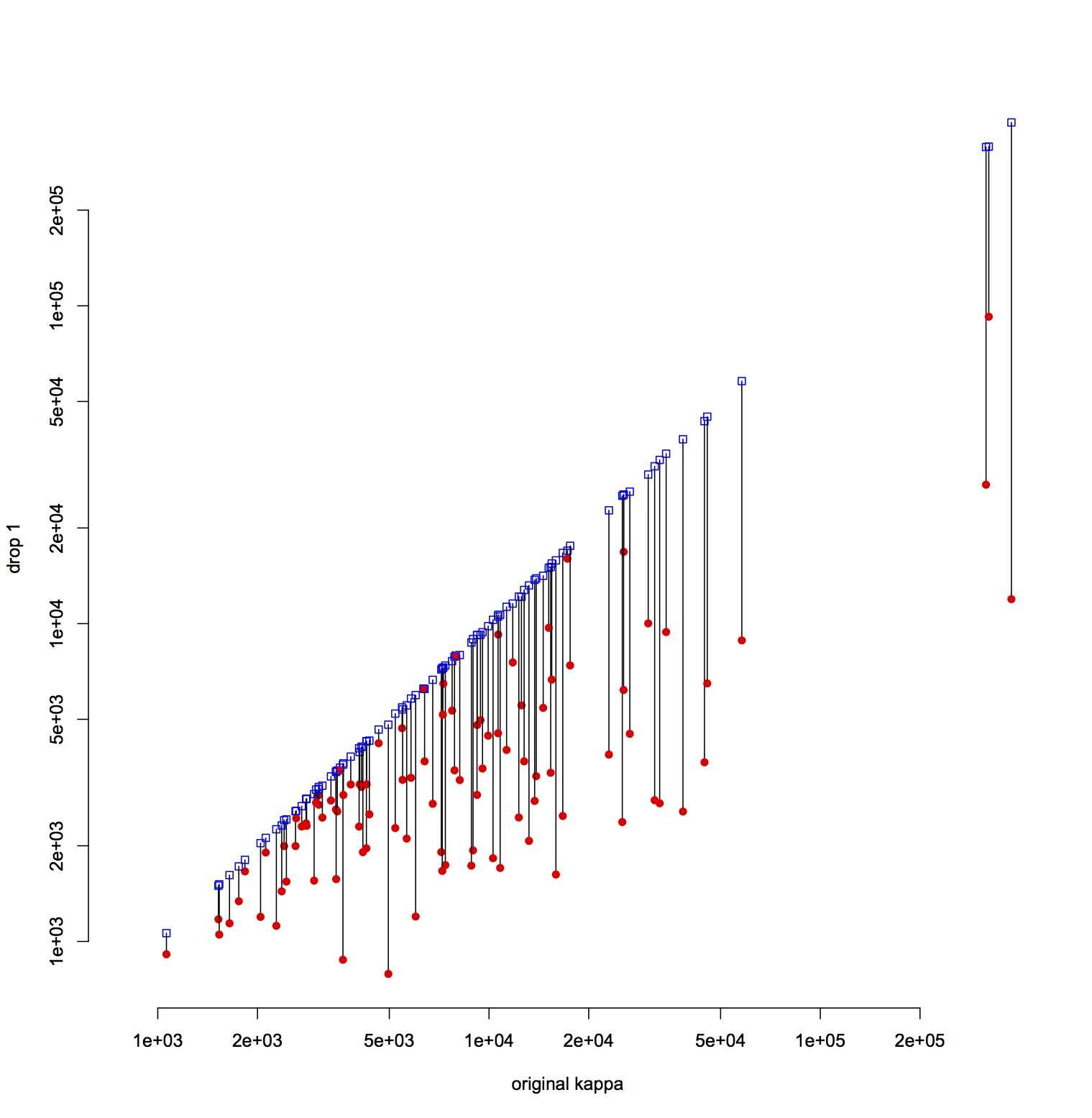
The matrix is

has two eigenvalues and and the condition number of defined by the ratio of the largest eigenvalues to the smallest eigenvalues is where is the big O notation that describes the limit behavior of a function. When we have tiny tip branch (very small ), the value of will be fairly large which indicates the matrix is more of ill-condition. For instance, with , while with , we have . The problem becomes serious as is very closed to zero as we have a matrix of two almost identical columns/rows which makes a singular matrix with . In general, for a tree of arbitrary taxa that includes a clade of sub tree described in this case, we will face the ill-condition problem. We can define a measure using the fraction to quantify the ill condition of the tree where is the tree height, is the smallest tip length and is the branch lengths from the root to the most recent common ancestor of the tips with smallest tip lengths. When approaches to zero, the fraction approaches to zero. In this case, as two columns/rows in the matrix are almost the same, the matrix will suffer an ill condition of order . We use a simulated 100 taxa birth-death tree. For a good tree (), = 1.45, while for the bad tree () , . From here the tree condition can be seen of order .

*Drop the shortest tip can lead a better kappa*

We will show that the new tree obtained from dropping the shortest tip of the original tree has a better (lower) kappa. Let be the by strictly ultrametric matrix and eigenvalues of . Let be the matrix obtained by dropping the shortest tip from the tree. is still an strictly ultrametic matrix of size by with eigenvalues . By a special case of the Cauchy’s interlacing theorem ( Ch. 10.1 in Parlett 1980), we have . The condition number, defined as the ratio of the largest eigenvalue to the smallest eigenvalue are computed as and for and , respectively. From some algebra calculation, we have . This implies that . From above, we can conclude that a lower can be achieved by dropping the smallest tips.

*A simulation on comparing the two condition numbers*

  
In the above plot, the x-axis is the original kappa, and the lines connect points corresponding to that given tree, but with one taxon removed: the blue square is removing a taxon at random, and the red dot removing the taxon with the shortest branch tip length. Either removal tends to help, but removing the taxon with the shortest branch helps more.

*References*

Nabben R. and Varga R.S. (1994) A linear algebra proof that the inverse of a strictly ultrametric matrix is a strictly diagonally dominant Stieltjes matrix. SIAM J. Matrix Anal. Appl. 15(1) 107-113.

Parlett B. (1980) The symmetric eigenvalue problem. Prentice-Hall, Englewood Cliffs, N.J.

Roger A.H. and Charles R.J. (1985) Matrix Analysis, Cambridge University Press. The Press Syndicates if the University of Cambridge ISBN 0-521-38632-2