How much does convergence matter?

Tommy Gibson

5/11/2021

Motivation

In optimizing transition rate parameters, it is unclear how much parameter estimates and the value of the objective function change with added iterations. Results up to this point have been obtained using one arbitrary set of initial values for parameters. Because the optimization algorithm we are using is a local solver, we should probably perform some sort of grid search over initial values in an effort to find the global minimum of the objective function. If parameter estimates and the objective function don't change much after some small or modest number of iterations, it may present an alternative to letting the optimizer run to convergence, which takes 10+ hours depending on the stopping criteria.

Details on the loss function, transition rate paramters, and constraints can be found in past documents. Here we try to numerically and graphically summarize how optimization output changes with added iterations. The optimization function accepts a value `maxiter`, for which the algorithm stops after `maxiter` iterations. We use values of 500, 1000, 3000, 5000, 12000, and 20000 (which stopped at 18881 because it converged).

Objective function

We don't know how steep the objective function's surface is towards a minimum. Below is

- A table of the value of the objective function after each number of iterations
- A plot showing how quickly the objective function is steered towards its local minimum.

Note that the value at 0 iterations is the value of the objective function at the initial values.

Table 1: Value of objective function for varying levels of maxiter

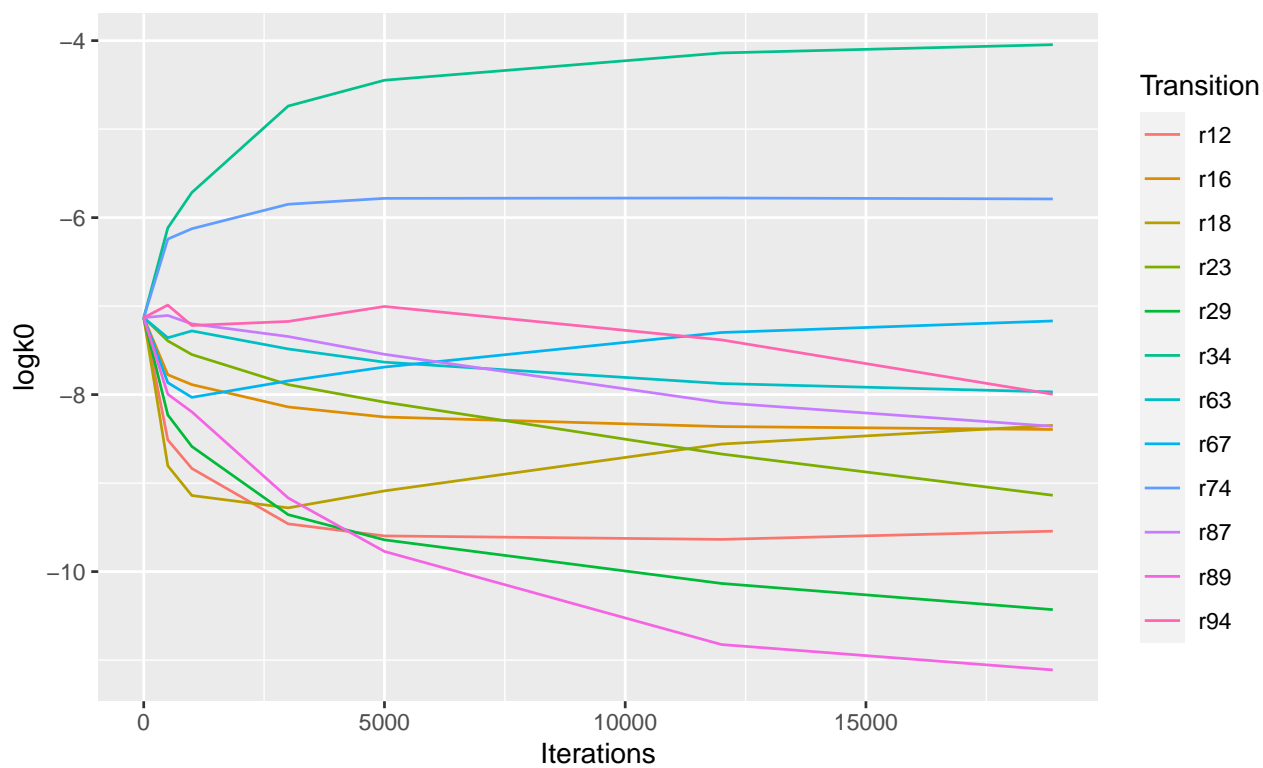| Iterations	Objective
0	1.8221274
500	0.3981999
1000	0.3803837
3000	0.3595116
5000	0.3529327
12000	0.3460737
18881	0.3441428

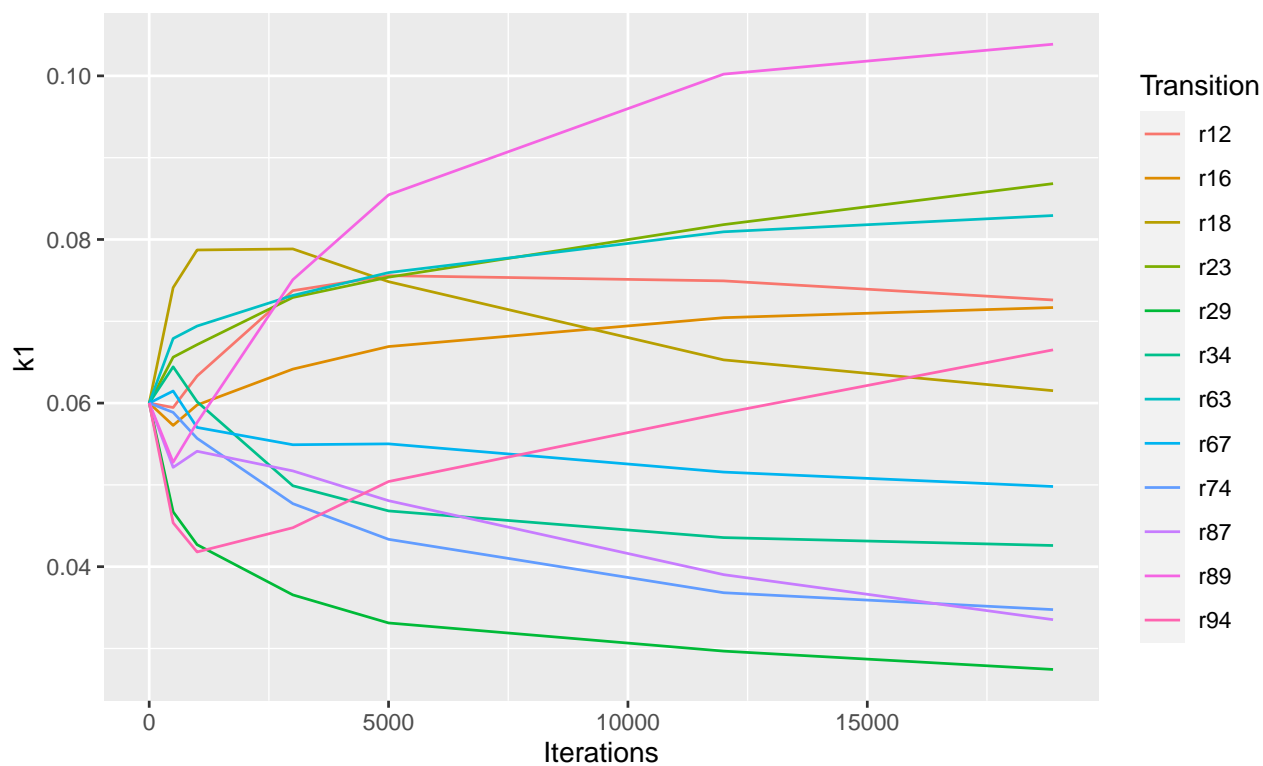Value of objective function for varying levels of maxiter
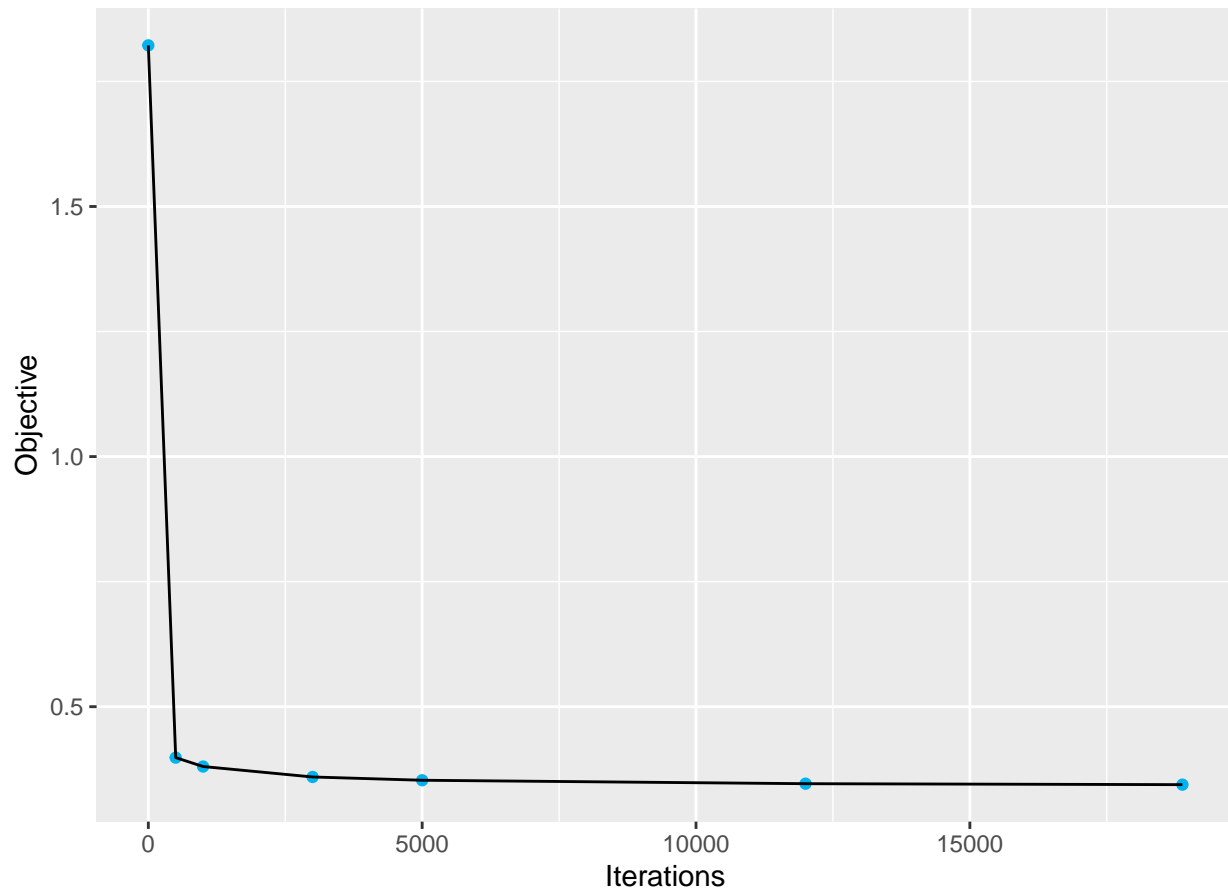
Parameters

Now we assess how much parameters change with varying levels of `maxiter`.

Parameter log(k0) by value of maxiter

Parameter k1 by value of maxiter

Transitions

We can calculate the transitions $r_{ij}(a)$ as functions of $\log(k_0)$ and k_1 for each value of `maxiter` to see how much the changes in parameter values actually affect the transition rates.

First, a plot of all transitions with the y-axis ranging from 0 to 1 in each plot.

Transition probabilities by maxiter
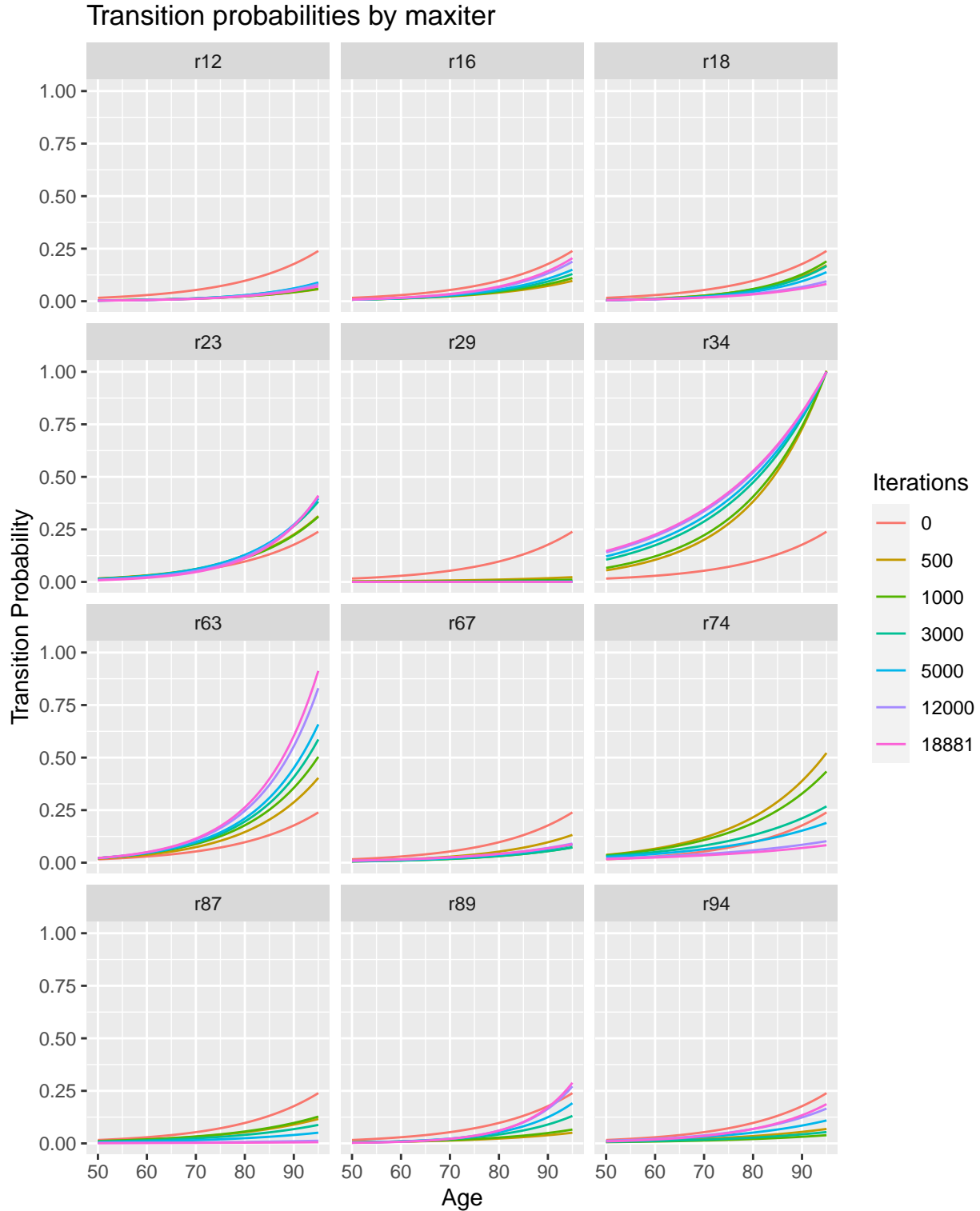
Now, a plot of all transitions where the y-axes are specific to each transition.

Transition probabilities by maxiter