

Application of **lldmn** to Human Microbiome Data

Peng Yu and Chad A. Shaw

We here provide an example on how to use our R package **lldmn** to analyze four human microbiome data set available from the R package **HMP**.

First, we load the required R packages.

```
suppressPackageStartupMessages(library(HMP))
suppressPackageStartupMessages(library(lldmn))
library(parallel)
```

Then, we make a function that returns the number of columns (K) and the corresponding Dirichlet-multinomial log-likelihood function (`loglikfun`) given the count data set (`count_table`).

```
dmnloglikfun_gen = function(count_table) {
  K = ncol(count_table)
  nsamples = nrow(count_table)
  list(K = K, loglikfun = function(x) {
    log_odd_ratio = x[seq_len(K - 1)]
    psi = x[[K]]

    p = (function(x) {
      x/sum(x)
    })(c(exp(log_odd_ratio), 1))
    sum(apply(count_table, 1, function(n) {
      lldmn(n, p, psi)
    })))
  })
}
```

We call the function `optim()` to calculate the maximum likelihood estimate of the dispersion parameter for each data set (`saliva`, `throat`, `tongue` and `tonsils`).

```
result = mclapply(list(saliva = saliva, throat = throat, tongue = tongue,
  tonsils = tonsils), function(count_table) {
  with(dmnloglikfun_gen(count_table), list(K = K, par = optim(par = rep(0,
```

```

    K), fn = loglikfun, lower = c(rep(-Inf, K - 1), 0), method = "L-BFGS-B",
    control = list(fnscale = -1, ndeps = rep(1e-07, K), maxit = .Machine$integer.max,
    factr = 0, lmm = 5 * K))$par))
}, mc.cores = parallel::detectCores())

```

Finally, we get the same estimates of the dispersion parameter ψ as shown in Table 1 of the main paper.

```

psi = lapply(result, function(x) {
  x$par[[x$K]]
})
psi

## $saliva
## [1] 0.003892
##
## $throat
## [1] 0.006394
##
## $tongue
## [1] 0.008016
##
## $tonsils
## [1] 0.01039

```

The proportion parameter \mathbf{p} (in this case representing taxa) can also be extracted as shown below.

```

prop = lapply(result, function(x) {
  x = c(exp(x$par[-x$K]), 1)
  x/sum(x)
})
prop

## $saliva
## [1] 0.185760 0.142173 0.116380 0.086501 0.064117 0.053694 0.041506
## [8] 0.033909 0.027122 0.023399 0.020321 0.018391 0.016307 0.015235
## [15] 0.013517 0.012345 0.011163 0.010372 0.009790 0.009081 0.088916
##
## $throat
## [1] 0.355399 0.127227 0.089583 0.067307 0.052411 0.043051 0.034516
## [8] 0.029562 0.024704 0.020709 0.017987 0.015666 0.013836 0.012327
## [15] 0.010828 0.009489 0.008258 0.007594 0.007255 0.006728 0.045564
##
## $tongue
## [1] 0.278580 0.141339 0.110892 0.081945 0.063612 0.050055 0.037374

```

```
## [8] 0.031572 0.025624 0.021515 0.018251 0.016566 0.014728 0.013327
## [15] 0.011372 0.010198 0.009025 0.008120 0.007596 0.007197 0.041112
##
## $tonsils
## [1] 0.367145 0.142237 0.086104 0.058657 0.045233 0.037113 0.031652
## [8] 0.028288 0.025229 0.020472 0.017351 0.015622 0.013802 0.012310
## [15] 0.011015 0.010026 0.009309 0.008726 0.008200 0.007560 0.043948
```