

Analysis of Junjian's Data

Tianyi Liu

10/05/2021

Seed Control

```
set.seed(100)
```

Read Data

We begin with reading in all the bulk, base, and proportion matrices:

```
## directory names
base_mtx_dir <- "./Real_data/Junjian_01/base_mtx"
bulk_dir <- "./Real_data/Junjian_01/bulk"
proportions_dir <- "./Real_data/Junjian_01/proportions"
output_dir <- "./Real_data/Junjian_01/output"

## all the base matrices
base_ABIS_logFC_a2s <- read_csv(paste(base_mtx_dir, "ABIS_logFC_a2s.csv", sep="/")) %>% rename(gene=X1)
base_immunoStates_logFC_a2s <- read_csv(paste(base_mtx_dir, "immunoStates_logFC_a2s.csv", sep="/")) %>%
base_LM22_logFC_a2s <- read_csv(paste(base_mtx_dir, "LM22_logFC_a2s.csv", sep="/")) %>% rename(gene=X1)
base_tianyi_logFC_a2s <- read_csv(paste(base_mtx_dir, "tianyi_logFC_a2s.csv", sep="/")) %>% rename(gene=X1)

## all the bulk matrices
bulk_markers_sd <- read_tsv(paste(bulk_dir, "bulk_markers_sd.txt", sep="/")) %>% rename(gene = GeneSymbol)
bulk_markers_sd_q <- read_tsv(paste(bulk_dir, "bulk_markers_sd_q.txt", sep="/")) %>% rename(gene = GeneSymbol)
bulk_nonmarkers_sd <- read_tsv(paste(bulk_dir, "bulk_nonmkrs_sd.txt", sep="/")) %>% rename(gene = GeneSymbol)
bulk_nonmarkers_sd_q <- read_tsv(paste(bulk_dir, "bulk_nonmkrs_sd_q.txt", sep="/")) %>% rename(gene = GeneSymbol)
bulk_Z0201 <- read_tsv(paste(bulk_dir, "Z0201_bulkCIBERSORT_a2s.txt", sep="/")) %>% rename(gene = GeneSymbol)

## all the proportion matrices
prop_tianyi <- read_csv(paste(proportions_dir, "CIBERSORTx_Job144_Results_tianyi.csv", sep="/"))
prop_lm22 <- read_csv(paste(proportions_dir, "CIBERSORTx_Job145_Results_lm22.csv", sep="/"))
prop_immunoStates <- read_csv(paste(proportions_dir, "CIBERSORTx_Job146_Results_immuno.csv", sep="/"))
prop_abis <- read_csv(paste(proportions_dir, "CIBERSORTx_Job147_Results_abis.csv", sep="/"))
```

Constants

We identify some constants from the data that are applicable to each of the matrices:

```
n_cell_types <- 6
n_subjects <- 51
```

Data Pre-processing

Next, we write a series of helper functions to obtain matrices for deconvolution:

```

get_bulk <- function(bulk_df) {
  # subset the expression and get gene and subject names
  gene_names = bulk_df$gene
  bulk_exp_df = bulk_df %>%
    select(2:(1 + n_subjects))
  subject_names = colnames(bulk_exp_df) %>%
    substr(2, 5)

  # assemble the matrix
  mat = data.matrix(bulk_exp_df)
  rownames(mat) = gene_names
  colnames(mat) = subject_names
  return(mat)
}

#get_bulk(bulk_markers_sd)

get_base <- function(base_df) {
  gene_names = base_df$gene
  base_exp_df = base_df %>%
    select(2:(1 + n_cell_types))
  cell_type_names = colnames(base_exp_df)

  mat = data.matrix(base_exp_df)
  rownames(mat) = gene_names
  colnames(mat) = cell_type_names
  return(mat)
}

#get_base(base_ABIS_logFC_a2s)

get_prop <- function(prop_df) {
  subject_names = prop_df$Mixture
  prop_real_df = prop_df %>%
    select(2:(1 + n_cell_types))
  cell_type_names = colnames(prop_real_df)

  mat = data.matrix(prop_real_df) %>% t
  rownames(mat) = cell_type_names
  colnames(mat) = subject_names
  return(mat)
}

#get_prop(prop_abis)

```

With the new functions we get all the 5 bulk matrices prepared:

```

bulk_markers_mat <- get_bulk(bulk_markers_sd)
bulk_markers_q_mat <- get_bulk(bulk_markers_sd_q)
bulk_nonmkrs_mat <- get_bulk(bulk_nonmarkers_sd)
bulk_nonmkrs_q_mat <- get_bulk(bulk_nonmarkers_sd_q)
bulk_Z0201_mat <- get_bulk(bulk_Z0201)

```

We also define some tuning parameters:

```
v_alpha = 2^seq(4, 8, 1)
v_xi = 2^-6
v_beta = 2^seq(2, 10, 2)
```

Main Helper Functions

We will use the following function to run the algorithm with 5 initial values given a set of tuning parameters and find the best run by the average column-wise correlations between the calculated and the truth:

```
## helper function to run the algorithm given a set of tuning parameter values and a pair of base and bulk matrices
## choose the best run as results
run_PSMF <- function(bulk, base, prop, alpha, xi, beta, k=5) {

  ## best avg cor
  best_cor = 0

  ## best solution
  best_sol = NULL

  ## get the gene names of base and bulk matrices
  x = rownames(base)
  y = rownames(bulk)

  ## genes that are in base but not in bulk matrix
  base_x = setdiff(x, y)

  ## genes that are in bulk but not base matrix
  bulk_y = setdiff(y, x)

  ## find the coordinates of this gene
  base_x_coords = which(x %in% base_x)
  bulk_y_coords = which(y %in% bulk_y)

  ## remove those genes from those coordinates
  GO_upper = base[-base_x_coords, ]
  M = bulk[-bulk_y_coords, ]

  ## get the number of unknown genes
  d_len = nrow(M) - nrow(GO_upper)

  ## fill GO
  mean_G = mean(base)
  GO_lower = matrix(0L, d_len, n_cell_types)
  GO = rbind(GO_upper, GO_lower)

  ## other parameters
  n_genes = nrow(M)
  n_known_genes = nrow(GO_upper)
  n_good_cell_types = n_cell_types

  ## run the algorithm with k different initial values and try to find the one with the best correlation
  for (i in 1:k) {
    G_init_lower = matrix(rchisq(d_len * n_cell_types, mean_G),
                          d_len,
```

```

n_cell_types)

G_init = rbind(G0_upper, G_init_lower)
C_init = get_C(n_cell_types, n_subjects)

solve_tmp = PSMF_solve(
  M = M,
  G_0 = G0,
  G_init = G_init,
  C_init = C_init,
  n_markers = n_known_genes,
  n_good_cell_types = n_good_cell_types,
  alpha = alpha,
  xi = xi,
  beta = beta
)

C_hat_tmp = solve_tmp$C_hat

## calculate the average correlation per cell type
# cor_tmp = colwise_cor(t(C_hat_tmp), t(prop))

## calculate the average correlation per sample
cor_tmp = colwise_cor(C_hat_tmp, prop)

if (is.na(cor_tmp)) {
  print("Unable to calculate correlations due to no variation in sample/cell type proportion estimation")
  cor_tmp = 0
}

if (cor_tmp > best_cor) {
  best_cor = cor_tmp
  best_sol = solve_tmp
}

}

if (best_cor == 0) {
  print("None of the solutions has positive correlation with the truth!")
  res = list(cor = 0, sol = NULL)
  return(res)
} else {
  #print(paste0("Best correlation = ", round(best_cor, 3)))
  res = list(cor = best_cor, sol = best_sol)
  return(res)
}

}

```

Then, we write another upper-level function to control the input of tuning parameters:

```

tune_PSMF <- function(v_alpha, v_xi, v_beta, ...) {

  ## the best correlation

```

```

best_cor = 0

## the best solution
best_sol = NULL

for (i in 1:length(v_alpha)) {
  for (j in 1:length(v_xi)) {
    for (k in 1:length(v_beta)) {

      ## get the tuning parameters
      alpha = v_alpha[i]
      xi = v_xi[j]
      beta = v_beta[k]

      ## run the algorithm with the above parameters
      solve = run_PSMF(..., alpha=alpha, xi=xi, beta=beta)

      ## collect the results
      cor_tmp = solve$cor

      ## replace the best result if the result improves
      if (cor_tmp > best_cor) {
        best_cor = cor_tmp
        best_sol = solve
      }

    }
  }
}

if (best_cor == 0) {
  print("Overall, none of the solutions has positive correlation with the truth!")
  res = list(cor = 0, sol = NULL, y_hat = NULL)
  return(res)
} else {
  print(paste0("Best correlation = ", round(best_cor, 3)))
  res = list(cor = best_cor, sol = best_sol)
  return(res)
}
}

```

ABIS

We run the deconvolution algorithm using the ABIS matrix as base matrix on 5 different bulk matrices

```

## get the abis base and prop matrix
base_abis_mat <- get_base(base_ABIS_logFC_a2s)
prop_abis_mat <- get_prop(prop_abis)

## run deconvolution on all bulks
ptm <- proc.time()
abis_tune_1 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,

```

```

        v_beta=v_beta,
        bulk=bulk_markers_mat,
        base=base_abis_mat,
        prop=prop_abis_mat)

## [1] "Best correlation = 0.724"

## save the estimated bulk
abis_y_hat1 <- abis_tune_1$sol$sol$G_hat %*% abis_tune_1$sol$sol$C_hat
write_csv(as_tibble(abis_y_hat1), paste(output_dir, "abis_all_markers.csv", sep="/"))

print("Completed the first bulk matrix.")

## [1] "Completed the first bulk matrix."
abis_tune_2 <- tune_PSMF(v_alpha=v_alpha,
        v_xi=v_xi,
        v_beta=v_beta,
        bulk=bulk_markers_q_mat,
        base=base_abis_mat,
        prop=prop_abis_mat)

## [1] "Best correlation = 0.567"

abis_y_hat2 <- abis_tune_2$sol$sol$G_hat %*% abis_tune_2$sol$sol$C_hat
write_csv(as_tibble(abis_y_hat2), paste(output_dir, "abis_interq_markers.csv", sep="/"))

print("Completed the second bulk matrix.")

## [1] "Completed the second bulk matrix."
abis_tune_3 <- tune_PSMF(v_alpha=v_alpha,
        v_xi=v_xi,
        v_beta=v_beta,
        bulk=bulk_nonmkrs_mat,
        base=base_abis_mat,
        prop=prop_abis_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.145"

abis_y_hat3 <- abis_tune_3$sol$sol$G_hat %*% abis_tune_3$sol$sol$C_hat
write_csv(as_tibble(abis_y_hat3), paste(output_dir, "abis_all_nonmkrs.csv", sep="/"))

print("Completed the third bulk matrix.")

## [1] "Completed the third bulk matrix."

## This is the estimated bulk we wanted
abis_tune_4 <- tune_PSMF(v_alpha=v_alpha,
        v_xi=v_xi,
        v_beta=v_beta,
        bulk=bulk_nonmkrs_q_mat,
        base=base_abis_mat,
        prop=prop_abis_mat)

```

```
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.218"

abis_y_hat4 <- abis_tune_4$sol$sol$G_hat %*% abis_tune_4$sol$sol$C_hat
write_csv(as_tibble(abis_y_hat4), paste(output_dir, "abis_interq_nonmkrs.csv", sep="/"))

print("Completed the fourth bulk matrix.")

## [1] "Completed the fourth bulk matrix."

abis_tune_5 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_Z0201_mat,
                        base=base_abis_mat,
                        prop=prop_abis_mat)

## [1] "Best correlation = 0.723"

abis_y_hat5 <- abis_tune_5$sol$sol$G_hat %*% abis_tune_5$sol$sol$C_hat
write_csv(as_tibble(abis_y_hat5), paste(output_dir, "abis_interq_Z0201.csv", sep="/"))

print("Completed the fifth bulk matrix.")

## [1] "Completed the fifth bulk matrix."

proc.time() - ptm

##      user  system elapsed
## 11.897   1.871  14.008
```

immonuoStates

```
## get the immunoStates base and prop matrix
base_immuno_mat <- get_base(base_immunoStates_logFC_a2s)
prop_immuno_mat <- get_prop(prop_immunoStates)

## run deconvolution on all bulks
ptm <- proc.time()
immuno_tune_1 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_markers_mat,
                          base=base_immuno_mat,
                          prop=prop_immuno_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.82"

## save the estimated bulk
if (immuno_tune_1$cor > 0) {
  immuno_y_hat1 <- immuno_tune_1$sol$sol$G_hat %*% immuno_tune_1$sol$sol$C_hat
  write_csv(as_tibble(immuno_y_hat1), paste(output_dir, "immuno_all_markers.csv", sep="/"))
}
```

```

print("Completed the first bulk matrix.")

## [1] "Completed the first bulk matrix."
immuno_tune_2 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_markers_q_mat,
                          base=base_immuno_mat,
                          prop=prop_immuno_mat)

## [1] "Best correlation = 0.247"
if (immuno_tune_2$cor > 0) {
  immuno_y_hat2 <- immuno_tune_2$sol$sol$G_hat %*% immuno_tune_2$sol$sol$C_hat
  write_csv(as_tibble(immuno_y_hat2), paste(output_dir, "immuno_interq_markers.csv", sep="/"))
}

print("Completed the second bulk matrix.")

## [1] "Completed the second bulk matrix."
immuno_tune_3 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_nonmkrs_mat,
                          base=base_immuno_mat,
                          prop=prop_immuno_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.168"
if (immuno_tune_3$cor > 0) {
  immuno_y_hat3 <- immuno_tune_3$sol$sol$G_hat %*% immuno_tune_3$sol$sol$C_hat
  write_csv(as_tibble(immuno_y_hat3), paste(output_dir, "immuno_all_nonmkrs.csv", sep="/"))
}

print("Completed the third bulk matrix.")

## [1] "Completed the third bulk matrix."
## This is the estimated bulk we wanted
immuno_tune_4 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_nonmkrs_q_mat,
                          base=base_immuno_mat,
                          prop=prop_immuno_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.199"
if (immuno_tune_4$cor > 0) {
  immuno_y_hat4 <- immuno_tune_4$sol$sol$G_hat %*% immuno_tune_4$sol$sol$C_hat
  write_csv(as_tibble(immuno_y_hat4), paste(output_dir, "immuno_interq_nonmkrs.csv", sep="/"))
}

```



```

}

print("Completed the fourth bulk matrix.")

## [1] "Completed the fourth bulk matrix."
immuno_tune_5 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_Z0201_mat,
                          base=base_immuno_mat,
                          prop=prop_immuno_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.63"

if (immuno_tune_5$cor > 0) {
  immuno_y_hat5 <- immuno_tune_5$sol$sol$G_hat %*% immuno_tune_5$sol$sol$C_hat
  write_csv(as_tibble(immuno_y_hat5), paste(output_dir, "immuno_interq_Z0201.csv", sep="/"))
}

print("Completed the fifth bulk matrix.")

## [1] "Completed the fifth bulk matrix."

proc.time() - ptm

##      user  system elapsed
## 116.255   20.930  138.944

```

LM22

```

## get the abis base and prop matrix
base_lm22_mat <- get_base(base_LM22_logFC_a2s)
prop_lm22_mat <- get_prop(prop_lm22)

## run deconvolution on all bulks
ptm <- proc.time()
lm22_tune_1 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_markers_mat,
                        base=base_lm22_mat,
                        prop=prop_lm22_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Overall, none of the solutions has positive correlation with the truth!"

## save the estimated bulk
if (lm22_tune_1$cor > 0) {
  lm22_y_hat1 <- lm22_tune_1$sol$sol$G_hat %*% lm22_tune_1$sol$sol$C_hat
  write_csv(as_tibble(lm22_y_hat1), paste(output_dir, "lm22_all_markers.csv", sep="/"))
}

print("Completed the first bulk matrix.")

## [1] "Completed the first bulk matrix."

lm22_tune_2 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_markers_q_mat,
                        base=base_lm22_mat,
                        prop=prop_lm22_mat)

## [1] "Best correlation = 0.197"

if (lm22_tune_2$cor > 0) {
  lm22_y_hat2 <- lm22_tune_2$sol$sol$G_hat %*% lm22_tune_2$sol$sol$C_hat
  write_csv(as_tibble(lm22_y_hat2), paste(output_dir, "lm22_interq_markers.csv", sep="/"))
}

print("Completed the second bulk matrix.")

## [1] "Completed the second bulk matrix."

lm22_tune_3 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_nonmkrs_mat,
                        base=base_lm22_mat,
                        prop=prop_lm22_mat)

## [1] "Best correlation = 0.183"

```

```

if (lm22_tune_3$cor > 0) {
  lm22_y_hat3 <- lm22_tune_3$sol$sol$G_hat %*% lm22_tune_3$sol$sol$C_hat
  write_csv(as_tibble(lm22_y_hat3), paste(output_dir, "lm22_all_nonmkrs.csv", sep="/"))
}

```

```

print("Completed the third bulk matrix.")

```

```

## [1] "Completed the third bulk matrix."

```

```

## This is the estimated bulk we wanted
lm22_tune_4 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_nonmkrs_q_mat,
                        base=base_lm22_mat,
                        prop=prop_lm22_mat)

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "Best correlation = 0.159"

```

```

if (lm22_tune_4$cor > 0) {
  lm22_y_hat4 <- lm22_tune_4$sol$sol$G_hat %*% lm22_tune_4$sol$sol$C_hat
  write_csv(as_tibble(lm22_y_hat4), paste(output_dir, "lm22_interq_nonmkrs.csv", sep="/"))
}

```

```

print("Completed the fourth bulk matrix.")

```

```

## [1] "Completed the fourth bulk matrix."

```

```

lm22_tune_5 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_Z0201_mat,
                        base=base_lm22_mat,
                        prop=prop_lm22_mat)

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"

```

```
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Overall, none of the solutions has positive correlation with the truth!"

if (lm22_tune_5$cor > 0) {
  lm22_y_hat5 <- lm22_tune_5$sol$sol$G_hat %*% lm22_tune_5$sol$sol$C_hat
  write_csv(as_tibble(lm22_y_hat5), paste(output_dir, "lm22_interq_Z0201.csv", sep="/"))
}

print("Completed the fifth bulk matrix.")

## [1] "Completed the fifth bulk matrix."

proc.time() - ptm

##      user  system elapsed
##    2.227    0.198    2.464
```

tianyi

```
## get the abis base and prop matrix
base_tianyi_mat <- get_base(base_tianyi_logFC_a2s)
prop_tianyi_mat <- get_prop(prop_tianyi)

## get the genes in the base matrix
gene_tianyi <- rownames(base_tianyi_mat)

## run deconvolution on all bulks
ptm <- proc.time()
tianyi_tune_1 <- tune_PSMF(v_alpha=v_alpha,
                          v_xi=v_xi,
                          v_beta=v_beta,
                          bulk=bulk_markers_mat,
                          base=base_tianyi_mat,
                          prop=prop_tianyi_mat)
```

```
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
```

```

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Overall, none of the solutions has positive correlation with the truth!"

## save the estimated bulk
if (tiany_i_tune_1$cor > 0) {
  tiany_i_y_hat1 <- tiany_i_tune_1$sol$sol$G_hat %*% tiany_i_tune_1$sol$sol$C_hat
  write_csv(as_tibble(tiany_i_y_hat1), paste(output_dir, "tiany_i_all_markers.csv", sep="/"))
}

print("Completed the first bulk matrix.")

## [1] "Completed the first bulk matrix."

tiany_i_tune_2 <- tune_PSMF(v_alpha=v_alpha,
                           v_xi=v_xi,
                           v_beta=v_beta,
                           bulk=bulk_markers_q_mat,
                           base=base_tiany_i_mat,
                           prop=prop_tiany_i_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.289"

if (tiany_i_tune_2$cor > 0) {
  tiany_i_y_hat2 <- tiany_i_tune_2$sol$sol$G_hat %*% tiany_i_tune_2$sol$sol$C_hat
  write_csv(as_tibble(tiany_i_y_hat2), paste(output_dir, "tiany_i_interq_markers.csv", sep="/"))
}

print("Completed the second bulk matrix.")

## [1] "Completed the second bulk matrix."

```

```

tianyitune_3 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_nonmkrs_mat,
                        base=base_tianyi_mat,
                        prop=prop_tianyi_mat)

## [1] "Best correlation = 0.153"

if (tianyitune_3$cor > 0) {
  tianyi_y_hat3 <- tianyitune_3$sol$sol$G_hat %*% tianyitune_3$sol$sol$C_hat
  write_csv(as_tibble(tianyi_y_hat3), paste(output_dir, "tianyi_all_nonmkrs.csv", sep="/"))
}

print("Completed the third bulk matrix.")

## [1] "Completed the third bulk matrix."
## This is the estimated bulk we wanted
tianyitune_4 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_nonmkrs_q_mat,
                        base=base_tianyi_mat,
                        prop=prop_tianyi_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Best correlation = 0.138"

if (tianyitune_4$cor > 0) {
  tianyi_y_hat4 <- tianyitune_4$sol$sol$G_hat %*% tianyitune_4$sol$sol$C_hat
  write_csv(as_tibble(tianyi_y_hat4), paste(output_dir, "tianyi_interq_nonmkrs.csv", sep="/"))
}

print("Completed the fourth bulk matrix.")

## [1] "Completed the fourth bulk matrix."

tianyitune_5 <- tune_PSMF(v_alpha=v_alpha,
                        v_xi=v_xi,
                        v_beta=v_beta,
                        bulk=bulk_Z0201_mat,
                        base=base_tianyi_mat,
                        prop=prop_tianyi_mat)

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"

```

```

## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "None of the solutions has positive correlation with the truth!"
## [1] "Overall, none of the solutions has positive correlation with the truth!"

if (tianyi_tune_5$cor > 0) {
  tianyi_y_hat5 <- tianyi_tune_5$sol$sol$G_hat %*% tianyi_tune_5$sol$sol$C_hat
  write_csv(as_tibble(tianyi_y_hat5), paste(output_dir, "tianyi_interq_Z0201.csv", sep="/"))
}

print("Completed the fifth bulk matrix.")

## [1] "Completed the fifth bulk matrix."

proc.time() - ptm

##      user  system elapsed
##    7.596    1.176    8.910

```