

```
#####
```

```
### This program is written in R. It is fully executable as written in a R environment
```

```
#####
```

```
## Require packages and read in the the data collateral provided.
```

```
require(plyr)
require(reshape2)
setwd("/Users/sarpotd/Desktop/Coursera/Recommender Systems/week7/Assignment1/svd-assignment")
ratings <- read.csv("data/ratings.csv", header=FALSE)
colnames(ratings) <- c("user_id", "movie_id", "ratings")
users <- read.csv("data/users.csv", header=FALSE)
colnames(users) <- c("user_id", "user_name")
movies <- read.csv("data/movie-titles.csv", header=FALSE)
colnames(movies) <- c("movie_id", "movie_name")
```

```
## Data munging to change it to a form useful for manipulation
## user_ratings as ratings of users for movies in a matrix form.
```

```
temp_1 <- melt(ratings, id=c("user_id","movie_id"), measure="ratings")
temp_2 <- acast(temp_1, user_id ~ movie_id ~ variable)
user_ratings <- as.matrix(temp_2[, ,1])
```

```
## Backup copy of the user_ratings matrix before we do data munging on it.
```

```
user_ratings_rownames <- rownames(user_ratings)
user_ratings_colnames <- colnames(user_ratings)
user_ratings_bak <- user_ratings
user_ratings_bak[is.na(user_ratings)] = 0
```

```
### ITEM
```

```
item_mean_function <- function(user,item) {

    mean_calc <- function(c) (mean(na.omit(c)))
    item_ratings_mean <- as.matrix(apply(user_ratings,2,mean_calc))
    rownames(item_ratings_mean) <- user_ratings_colnames

    sub_mean <- function(c) (c - item_ratings_mean)
    temp_1 <- apply(user_ratings_bak,1,sub_mean)
    rownames(temp_1) <- user_ratings_colnames
```

```

colnames(temp_1) <- user_ratings_rownames
item_ratings_mean_centered <- t(temp_1)
item_ratings_mean_centered[is.na(user_ratings)] = 0

ratings_for_svd <- item_ratings_mean_centered
ratings_for_svd[is.na(user_ratings)] = 0

ratings_svd <- svd(ratings_for_svd, nu = 10, nv = 10)
sigma <- diag(ratings_svd$d, 10, 10)
u <- ratings_svd$u
v <- ratings_svd$v
rownames(u) <- user_ratings_rownames
rownames(v) <- user_ratings_colnames

u[user,] %*% sigma %*% v[item,] + item_ratings_mean[item,]
}

### GLOBAL MEAN (WORKING)
global_mean_function <- function(user, item) {

  global_mean <- mean(user_ratings, na.rm = TRUE)
  user_ratings_global_mean <- matrix(data = global_mean, nrow = nrow(user_ratings),
ncol = ncol(user_ratings))
  rownames(user_ratings_global_mean) <- user_ratings_rownames
  colnames(user_ratings_global_mean) <- user_ratings_colnames

  ratings_for_svd <- user_ratings_bak - user_ratings_global_mean
  ratings_for_svd[is.na(user_ratings)] = 0

  ratings_svd <- svd(ratings_for_svd, nu = 10, nv = 10)
  sigma <- diag(ratings_svd$d, 10, 10)
  u <- ratings_svd$u
  v <- ratings_svd$v
  rownames(u) <- user_ratings_rownames
  rownames(v) <- user_ratings_colnames

  u[user,] %*% sigma %*% v[item,] + global_mean
}

### USER_ITEM (WORKING)
user_item_mean_function <- function(user, item) {

  mean_calc <- function(c) (mean(na.omit(c)))
  item_ratings_mean <- as.matrix(apply(user_ratings, 2, mean_calc))
  rownames(item_ratings_mean) <- user_ratings_colnames

  sub_mean <- function(c) (c - item_ratings_mean)
  temp_1 <- apply(user_ratings_bak, 1, sub_mean)

```

```

rownames(temp_1) <- user_ratings_colnames
colnames(temp_1) <- user_ratings_rownames
item_ratings_mean_centered <- t(temp_1)
item_ratings_mean_centered[is.na(user_ratings)] = NA

user_ratings_mean <- as.matrix(apply(item_ratings_mean_centered,1,mean_calc))
rownames(user_ratings_mean) <- user_ratings_rownames

item_ratings_mean_centered[is.na(user_ratings)] = 0

sub_mean <- function(c) (c - user_ratings_mean)
temp_1 <- apply(item_ratings_mean_centered,2,sub_mean)
rownames(temp_1) <- user_ratings_rownames
user_item_ratings_mean_centered <- temp_1
user_item_ratings_mean_centered[is.na(user_ratings)] = 0

ratings_for_svd <- user_item_ratings_mean_centered
ratings_for_svd[is.na(user_ratings)] = 0

ratings_svd <- svd(ratings_for_svd,nu =10, nv=10)
sigma <- diag(ratings_svd$d,10,10)
u <- ratings_svd$u
v <- ratings_svd$v
rownames(u) <- user_ratings_rownames
rownames(v) <- user_ratings_colnames

u[user,]%*%sigma%*%v[item,] + user_ratings_mean[user,] + item_ratings_mean[item,]
}

```

USER (WORKING)

```

user_mean_function <- function(user,item) {

  mean_calc <- function(c) (mean(na.omit(c)))
  user_ratings_mean <- as.matrix(apply(user_ratings,1,mean_calc))
  rownames(user_ratings_mean) <- user_ratings_rownames

  sub_mean <- function(c) (c - user_ratings_mean)
  temp_1 <- apply(user_ratings_bak,2,sub_mean)
  rownames(temp_1) <- user_ratings_rownames
  user_ratings_mean_centered <- temp_1
  user_ratings_mean_centered[is.na(user_ratings)] = 0

  ratings_for_svd <- user_ratings_mean_centered
  ratings_for_svd[is.na(user_ratings)] = 0

```

```

ratings_svd <- svd(ratings_for_svd,nu =10, nv=10)
sigma <- diag(ratings_svd$d,10,10)
u <- ratings_svd$u
v <- ratings_svd$v
rownames(u) <- user_ratings_rownames
rownames(v) <- user_ratings_colnames

u[user,]%*%sigma%*%v[item,] + user_ratings_mean[user,]
}

input_user_list <-
c(rep("5120",times=5),rep("926",times=5),rep("2387",times=5),rep("5394",times=5),rep("3435",times=5))
input_item_list <-
c("63","393","2164","8358","194","788","141","581","63","568","7443","9741","36658","9331",
,"134","808","141","752","3049","629","9741","601","105","3049","8358")
baseline <- c("global_mean","item_mean","user_mean","user_item_mean")

#mean_t <- "item_mean"

### Call the final_fun function on each of the user:movie combination iteratively.
for(mean_t in baseline) {
  if(mean_t == "global_mean") {
    i = as.integer(1)
    print("Global Mean")
    print(c(rep("-",times=100)))
    while (i <= length(input_user_list)) {
      pred <-
round(global_mean_function(input_user_list[i],input_item_list[i]), digits=4)
      print(input_user_list[i])
      print(input_item_list[i])
      print(pred)
      print(movies[movies
$movie_id==input_item_list[i],"movie_name"],max.levels=0)
      i = i+1
    }
  } else if(mean_t == "user_mean") {
    i = as.integer(1)
    print("User Mean")
    print(c(rep("-",times=100)))
    while (i <= length(input_user_list)) {
      pred <-
round(user_mean_function(input_user_list[i],input_item_list[i]), digits=4)
      print(input_user_list[i])
      print(input_item_list[i])
      print(pred)
      print(movies[movies
$movie_id==input_item_list[i],"movie_name"],max.levels=0)
      i = i+1
    }
  } else if(mean_t == "item_mean") {

```

```

        i = as.integer(1)
        print("Item Mean")
        print(c(rep("-",times=100)))
        while (i <= length(input_user_list)) {
            pred <-
round(item_mean_function(input_user_list[i],input_item_list[i]), digits=4)
            print(input_user_list[i])
            print(input_item_list[i])
            print(pred)
            print(movies[movies
$movie_id==input_item_list[i],"movie_name"],max.levels=0)
            i = i+1
        }
    } else if(mean_t == "user_item_mean") {
        i = as.integer(1)
        print("User Item Mean")
        print(c(rep("-",times=100)))
        while (i <= length(input_user_list)) {
            pred <-
round(user_item_mean_function(input_user_list[i],input_item_list[i]), digits=4)
            print(input_user_list[i])
            print(input_item_list[i])
            print(pred)
            print(movies[movies
$movie_id==input_item_list[i],"movie_name"],max.levels=0)
            i = i+1
        }
    }
}

```