

# Package ‘MatchGATE’

April 8, 2024

**Title** Estimate Group Average Treatment Effects with Matching

**Version** 0.0.10

**Description** Two novel matching-based methods for estimating group average treatment effects (GATEs). The `match_y1y0()` and `match_y1y0_bc()` functions are used for imputing the potential outcomes based on matching and bias-corrected matching techniques, respectively. The `EstGATE()` function is employed to estimate the GATE after imputing the potential outcomes.

**License** GPL-3

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Imports** locpol, stats

**NeedsCompilation** no

**Author** Zhaoqing Tian [aut, cre, com] (<<https://orcid.org/0009-0001-6786-0924>>),  
Peng Wu [aut, ths] (<<https://orcid.org/0000-0001-7154-8880>>),  
Yilin Chen [dct] (<<https://orcid.org/0009-0009-2418-1348>>)

**Maintainer** Zhaoqing Tian <tzqluck@163.com>

**Repository** CRAN

**Date/Publication** 2024-04-08 15:10:05 UTC

## R topics documented:

EstGATE . . . . .	2
match_y1y0 . . . . .	3
match_y1y0_bc . . . . .	4
<b>Index</b>	<b>6</b>

**Description**

When imputed values for  $Y^1$  and  $Y^0$  are available for each individual, we can use EstGATE to estimate the group average treatment effects (GATE) defined by

$$GATE(z) = E[Y^1 - Y^0 | Z = z]$$

for some for possible values  $z$  of  $Z$ .

**Usage**

```
EstGATE(Y1_Y0, Z, Zeval, h)
```

**Arguments**

<code>Y1_Y0</code>	A vector in which each element is a treatment effect for each individual.
<code>Z</code>	A subvector of the covariates $X$ , which is used to define the subgroup of interest.
<code>Zeval</code>	Vector of evaluation points of $Z$ .
<code>h</code>	A smoothing parameter, bandwidth.

**Value**

The value of the corresponding GATE at different evaluation points.

**Examples**

```
set.seed(691)
n <- 2000
X1 <- runif(n, -0.5, 0.5)
X2 <- rnorm(n, sd = 0.5)
X = cbind(X1, X2)
A = sample(c(0,1), n, TRUE)
Y0 <- X2 + X1*X2/2 + rnorm(n, sd = 0.25)
Y1 <- A * (2*X1^2) + X2 + X1*X2/2 + rnorm(n, sd = 0.25)
Y <- A * Y1 + (1-A)*Y0
res.match <- match_y1y0(X, A, Y, K = 5)
y1_y0 <- res.match$Y1 - res.match$Y0
Z <- X1
Zeval = seq(min(Z), max(Z), len = 101)
h <- 0.5 * n^(-1/5)
res <- EstGATE(Y1_Y0 = y1_y0, Z, Zeval, h = h)
plot(x = Zeval, y = 2*Zeval^2,
     type = "l", xlim = c(-0.6, 0.5),
     main = "Estimated value vs. true value",
```

```

xlab = "Zeval", ylab = "GATE",
col = "DeepPink", lwd = "2")
lines(x = res$Zeval, y = res$GATE,
      col="DarkTurquoise", lwd = "2")
legend('bottomleft', c("Estimated GATE", "True GATE"),
      col=c("DarkTurquoise", "DeepPink"),
      text.col=c("DarkTurquoise", "DeepPink"), cex = 0.8)

```

match\_y1y0

*Imputing Missing Potential Outcomes with Matching***Description**

Impute missing potential outcomes for each individual with matching.

**Usage**

```
match_y1y0(X, A, Y, K = 5, method = "euclidean")
```

**Arguments**

X	A matrix representing covariates, where each row represents the value of a different covariates for an individual.
A	A vector representing the treatment received by each individual.
Y	A vector representing the observed outcome for each individual.
K	When imputing missing potential outcomes, the average number of similar individuals are taken based on covariates similarity.
method	The distance measure to be used. It is a argument embed in <a href="#">dist</a> function.

**Details**

Here are the implementation details for the imputation processes. Denote  $\hat{Y}_i^0$  and  $\hat{Y}_i^1$  as the imputed potential outcomes for individual  $i$ . Without loss of generality, if  $A_i = 0$ , then  $\hat{Y}_i^0 = Y_i$ , and  $\hat{Y}_i^1$  is the average of outcomes for the  $K$  units that are the most similar to the individual  $i$ , i.e.,

$$\hat{Y}_i^1 = \frac{1}{K} \sum_{j \in \mathcal{J}_K(i)} Y_j,$$

where  $\mathcal{J}_K(i)$  represents the set of  $K$  matched individuals with  $A_j = 1$ , that are the closest to the individual  $i$  in terms of covariates similarity, and vice versa.

**Value**

Returns a matrix of completed matches, where each row is the imputed  $(Y^1, Y^0)$  for each individual.

**Examples**

```
n <- 100
p <- 2
X <- matrix(rnorm(n*p), ncol = p)
A <- sample(c(0,1), n, TRUE)
Y <- A * (2*X[,1]) + X[,2]^2 + rnorm(n)
match_y1y0(X = X, A = A, Y = Y, K = 5)
```

match\_y1y0\_bc

*Imputing Missing Potential Outcomes with Bias-Corrected Matching***Description**

Impute missing potential outcomes for each individual with bias-corrected matching.

**Usage**

```
match_y1y0_bc(X, A, Y, miu1.hat, miu0.hat, K = 5, method = "euclidean")
```

**Arguments**

X	A matrix representing covariates, where each row represents the value of a different covariates for an individual.
A	A vector representing the treatment received by each individual.
Y	A vector representing the observed outcome for each individual.
miu1.hat	The estimated outcome regression function for $Y^1$ .
miu0.hat	The estimated outcome regression function for $Y^0$ .
K	When imputing missing potential outcomes, the average number of similar individuals are taken based on covariates similarity.
method	The distance measure to be used. It is a argument embed in <a href="#">dist</a> function.

**Details**

Here are the implementation details for the imputation processes. Denote  $\hat{Y}_i^0$  and  $\hat{Y}_i^1$  as the imputed potential outcomes for individual  $i$ . For example, if  $A_i = 0$ , then  $\hat{Y}_i^0 = Y_i^0$ . However, for obtaining  $\hat{Y}_i^1$ , we require to introduce an outcome regression function  $\mu_1(X)$  for  $Y^1$ . Let  $\hat{\mu}_1(X)$  be the fitted value of  $\mu_1(X)$ , then  $\hat{Y}_i^1$  is defined as follows,

$$\hat{Y}_i^1 = \frac{1}{K} \sum_{j \in \mathcal{J}_K(i)} \{Y_j + \hat{\mu}_1(X_i) - \hat{\mu}_1(X_j)\},$$

where  $\mathcal{J}_K(i)$  represents the set of  $K$  matched individuals with  $A_j = 1$ , that are the closest to the individual  $i$  in terms of covariates similarity, and vice versa.

**Value**

Returns a matrix of completed matches, where each row is the imputed  $(Y^1, Y^0)$  for each individual.

**Examples**

```
n = 100
X1 <- runif(n, -0.5, 0.5)
X2 <- sample(c(0,1,2), n, TRUE)
X = cbind(X1, X2)
A = sample(c(0,1), n, TRUE)
Y = A * (2*X1) + X1 + X2^2 + rnorm(n)
miu1_hat <- cbind(1,X) %*% as.matrix(lm(Y ~ X, subset = A==1)$coef)
miu0_hat <- cbind(1,X) %*% as.matrix(lm(Y ~ X, subset = A==0)$coef)
match_y1y0_bc(X = X, A = A, Y = Y, miu1.hat = miu1_hat,
              miu0.hat = miu0_hat, K = 5)
```

# Index

dist, [3](#), [4](#)

EstGATE, [2](#)

match\_y1y0, [3](#)

match\_y1y0\_bc, [4](#)