

# Package ‘SSOSVM’

September 20, 2025

**Type** Package

**Title** Stream Suitable Online Support Vector Machines

**Version** 0.2.2

**Date** 2025-09-20

**Description** Soft-margin support vector machines (SVMs) are a common class of classification models. The training of SVMs usually requires that the data be available all at once in a single batch, however the Stochastic majorization-minimization (SMM) algorithm framework allows for the training of SVMs on streamed data instead Nguyen, Jones & McLachlan(2018)<doi:10.1007/s42081-018-0001-y>. This package utilizes the SMM framework to provide functions for training SVMs with hinge loss, squared-hinge loss, and logistic loss.

**License** GPL-3

**Encoding** UTF-8

**Imports** Rcpp (>= 0.12.13), mvtnorm

**LinkingTo** Rcpp, RcppArmadillo

**RoxygenNote** 6.1.1

**Suggests** testthat, knitr, rmarkdown, ggplot2, gganimate, gifski

**NeedsCompilation** yes

**Author** Andrew Thomas Jones [aut, cre],  
Hien Duy Nguyen [aut],  
Geoffrey J. McLachlan [aut]

**Maintainer** Andrew Thomas Jones <andrewthomasjones@gmail.com>

**Repository** CRAN

**Date/Publication** 2025-09-20 09:10:02 UTC

## Contents

|                       |   |
|-----------------------|---|
| generateSim . . . . . | 2 |
| Hinge . . . . .       | 2 |
| Logistic . . . . .    | 3 |
| SquareHinge . . . . . | 4 |
| SSOSVM . . . . .      | 5 |
| SVMFit . . . . .      | 5 |

---

|             |                             |
|-------------|-----------------------------|
| generateSim | <i>Generate Simulations</i> |
|-------------|-----------------------------|

---

**Description**

Generate simple simulations for testing of the algorithms.

**Usage**

generateSim(NN = 10^4, DELTA = 2, DIM = 2, seed = NULL)

**Arguments**

- |       |                                                              |
|-------|--------------------------------------------------------------|
| NN    | Number of observations. Default is 10^4                      |
| DELTA | Separation of three groups in standard errors. Default is 2. |
| DIM   | Number of dimensions in data. Default is 2.                  |
| seed  | Random seed if desired.                                      |

**Value**

A list containing:

- |      |                                             |
|------|---------------------------------------------|
| XX   | Coordinates of the simulated points.        |
| YY   | Cluster membership of the simulated points. |
| YMAT | YY and XX Combined as a single matrix.      |

**Examples**

```
#100 points of dimension 4.  
generateSim(NN=100, DELTA=2, DIM=4)
```

---

|       |              |
|-------|--------------|
| Hinge | <i>Hinge</i> |
|-------|--------------|

---

**Description**

Fit SVM with Hinge loss function.

**Usage**

Hinge(YMAT, DIM = 2L, EPSILON = 1e-05, returnAll = FALSE, rho = 1)

**Arguments**

|           |                                                                                                                                       |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------|
| YMAT      | Data. First column is -1 or 1 indicating the class of each observation. The remaining columns are the coordinates of the data points. |
| DIM       | Dimension of data. Default value is 2.                                                                                                |
| EPSILON   | Small perturbation value needed in calculation. Default value is 0.00001.                                                             |
| returnAll | Return all of theta values? Boolean with default value FALSE.                                                                         |
| rho       | Sensitivity factor to adjust the level of change in the SVM fit when a new observation is added. Default value 1.0                    |

**Value**

A list containing:

|            |                                                                                                          |
|------------|----------------------------------------------------------------------------------------------------------|
| THETA      | SVM fit parameters.                                                                                      |
| NN         | Number of observation points in YMAT.                                                                    |
| DIM        | Dimension of data.                                                                                       |
| THETA_list | THETA at each iteration (new point observed) as YMAT is fed into the algorithm one data point at a time. |
| OMEGA      | Intermediate value OMEGA at each iteration (new point observed).                                         |

**Examples**

```
YMAT <- generateSim(10^4)
h1<-Hinge(YMAT$YMAT,returnAll=TRUE)
```

---

Logistic

---

*Logistic Loss Function*


---

**Description**

Fit SVM with Logistic loss function.

**Usage**

```
Logistic(YMAT, DIM = 2L, EPSILON = 1e-05, returnAll = FALSE,
  rho = 1)
```

**Arguments**

|           |                                                                                                                                       |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------|
| YMAT      | Data. First column is -1 or 1 indicating the class of each observation. The remaining columns are the coordinates of the data points. |
| DIM       | Dimension of data. Default value is 2.                                                                                                |
| EPSILON   | Small perturbation value needed in calculation. Default value is 0.00001.                                                             |
| returnAll | Return all of theta values? Boolean with default value FALSE.                                                                         |
| rho       | Sensitivity factor to adjust the level of change in the SVM fit when a new observation is added. Default value 1.0                    |

**Value**

A list containing:

|            |                                                                                                          |
|------------|----------------------------------------------------------------------------------------------------------|
| THETA      | SVM fit parameters.                                                                                      |
| NN         | Number of observation points in YMAT.                                                                    |
| DIM        | Dimension of data.                                                                                       |
| THETA_list | THETA at each iteration (new point observed) as YMAT is fed into the algorithm one data point at a time. |
| CHI        | Intermediate value CHI at each iteration (new point observed).                                           |

**Examples**

```
YMAT <- generateSim(10^4)
l1<-Logistic(YMAT$YMAT,returnAll=TRUE)
```

---

|             |                     |
|-------------|---------------------|
| SquareHinge | <i>Square Hinge</i> |
|-------------|---------------------|

---

**Description**

Fit SVM with Square Hinge loss function.

**Usage**

```
SquareHinge(YMAT, DIM = 2L, EPSILON = 1e-05, returnAll = FALSE,
  rho = 1)
```

**Arguments**

|           |                                                                                                                                       |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------|
| YMAT      | Data. First column is -1 or 1 indicating the class of each observation. The remaining columns are the coordinates of the data points. |
| DIM       | Dimension of data. Default value is 2.                                                                                                |
| EPSILON   | Small perturbation value needed in calculation. Default value is 0.00001.                                                             |
| returnAll | Return all of theta values? Boolean with default value FALSE.                                                                         |
| rho       | Sensitivity factor to adjust the level of change in the SVM fit when a new observation is added. Default value 1.0                    |

**Value**

A list containing:

|            |                                                                                                          |
|------------|----------------------------------------------------------------------------------------------------------|
| THETA      | SVM fit parameters.                                                                                      |
| NN         | Number of observation points in YMAT.                                                                    |
| DIM        | Dimension of data.                                                                                       |
| THETA_list | THETA at each iteration (new point observed) as YMAT is fed into the algorithm one data point at a time. |
| PSI        | Intermediate value PSI at each iteration (new point observed).                                           |

**Examples**

```
YMAT <- generateSim(10^3,DIM=3)
sq1<-SquareHinge(YMAT$YMAT, DIM=3, returnAll=TRUE)
```

---

|        |                                                                                                                                                            |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SSOSVM | <i>SSOSVM: A package for online training of soft-margin support vector machines (SVMs) using the Stochastic majorization–minimization (SMM) algorithm.</i> |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------|

---

**Description**

The SSOSVM package allows for the online training of Soft-margin support vector machines (SVMs) using the Stochastic majorization–minimization (SMM) algorithm. SquareHinge,Hinge and Logistic The function generateSim can also be used to generate simple test sets.

**Author(s)**

Andrew T. Jones, Hien D. Nguyen, Geoffrey J. McLachlan

**References**

Hien D. Nguyen, Andrew T. Jones and Geoffrey J. McLachlan. (2018). Stream-suitable optimization algorithms for some soft-margin support vector machine variants, Japanese Journal of Statistics and Data Science, vol. 1, Issue 1, pp. 81-108.

---

|        |                            |
|--------|----------------------------|
| SVMFit | <i>SSOSVM Fit function</i> |
|--------|----------------------------|

---

**Description**

This is the primary function for uses to fit SVMs using this package.

**Usage**

```
SVMFit(YMAT, method = "logistic", EPSILON = 1e-05, returnAll = FALSE,
      rho = 1)
```

**Arguments**

|           |                                                                                                                                       |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------|
| YMAT      | Data. First column is -1 or 1 indicating the class of each observation. The remaining columns are the coordinates of the data points. |
| method    | Choice of function used in SVM. Choices are 'logistic', 'hinge' and 'square-Hinge'. Default value is 'logistic'                       |
| EPSILON   | Small perturbation value needed in calculation. Default value is 0.00001.                                                             |
| returnAll | Return all of theta values? Boolean with default value FALSE.                                                                         |
| rho       | Sensitivity factor to adjust the level of change in the SVM fit when a new observation is added. Default value 1.0                    |

**Value**

A list containing:

|                 |                                                                                                                |
|-----------------|----------------------------------------------------------------------------------------------------------------|
| THETA           | SVM fit parameters.                                                                                            |
| NN              | Number of observation points in YMAT.                                                                          |
| DIM             | Dimension of data.                                                                                             |
| THETA_list      | THETA at each iteration (new point observed) as YMAT is fed into the algorithm one data point at a time.       |
| PSI, OMEGA, CHI | Intermediate value for PSI, OMEGA, or CHI (depending on method choice) at each iteration (new point observed). |

**Examples**

```
Sim<- generateSim(10^4)
m1<-SVMFit(Sim$YMAT)
```

# Index

`generateSim`, [2](#)

`Hinge`, [2](#)

`Logistic`, [3](#)

`SquareHinge`, [4](#)

`SSOSVM`, [5](#)

`SSOSVM-package (SSOSVM)`, [5](#)

`SVMFit`, [5](#)