Problem: Find w that minimize function $f(w)=L(w)+R(w)$

parameter

d = number of dimension of w

a = bound

OPE algorithm

1. Initialize w_0 ($w_0[i]$ in (0,1) for i = 1...d)
2. set a = 0, b = 0
3. for t = 1...T

    3.1 choose random a or b with probability {0.5 , 0.5}

    then increase by 1

    3.2 $F_t = a*L(w_t)+b*R(w_t)$

    3.3 $F_t' = a*L'(w_t)+b*R'(w_t)$

    3.4 Find s_t that minimize $\langle F_t', s_t\rangle$ with respect to $\sum_{i=1}^{d} |(s_t[i])| < a$

    3.5 set $w_t = w_{t-1} + \frac{(s_{t-1} - w_{t-1})}{t}$

Detail:

3.3 R is non-differentiable at some point w, so at that points, we choose random ***right differential*** or ***left differential*** with probability {0.5, 0.5}

3.4 to find s_t , we choose s_t to be a vertex of domain $D = x \subset R^d \, such \, that \sum_{i=1}^{d} |(x[i])| < a$

because D is convex, and 3.4 problem is linear programing.

So one element of s_t is a or $-a$, other elements are zeros.

In GIST package, authors implemented loss functions :

Least : $L(w)=1/(2n)|Xw-y|^2$ n is number of sample

L2SVM : $L(w)=1/(2n)\sum_j max(0,1-y_j*x_j'*w)^2$

Logistic : $L(w)=1/n\sum_j \log(1+\exp(-y_j*x_j'*w))$