

USAGE

```
CasANOVA(formula,lambdas = seq(0.1,3,.1), eps = 10^{-7})
```

ARGUMENTS

| | |
|---------|---|
| formula | Symbolic formula as in call to <i>lm</i> . Factors should be of class "factor". May also contain numeric variables. |
| lambdas | Grid of penalty parameters at which solution is computed. All values must be > 0. Best will be chosen via BIC. |
| eps | Convergence criterion for the algorithm. |

VALUE

| | |
|--------------|---|
| coefficients | Estimated coefficients chosen by minimum BIC. |
| fitted | Fitted values from the chosen coefficients. |
| lambda | The best lambda chosen by BIC from the input grid points. |
| lambda.grid | The ordered values of penalty parameters used (lambdas). |
| BIC | Value of BIC for each penalty parameter. |
| df | Number of unique non-zero coefficient estimates for each penalty parameter. |
| init.coefs | The initial (unpenalized) coefficient estimates. |

NOTES

The parameterization used in this function is **NOT** the overparameterized ANOVA model with sum to zero constraints as in Bondell and Reich (2009).

The parameterization treats the last level of the factor as the baseline (which agrees with the function *lm* when using factors). So the coefficients represent the differences from baseline. Hence any zeros mean that those levels should be grouped together with the baseline category, while equalities at non-zero values indicate groups of coefficients that would exclude the baseline.

REFERENCE

Bondell, H. D. and Reich, B. J. (2009). Simultaneous factor selection and collapsing levels in ANOVA. *Biometrics* **65**, 169-177.

EXAMPLE

```
n = 100
x1 = sample(1:5,n,replace=T)
x2 = rnorm(n)
x3 = sample(1:6,n,replace=T)
x4 = sample(1:8,n,replace=T)
y = x1 + x2 + rnorm(n)
example_fit = CasANOVA(y~as.factor(x1)+x2+as.factor(x3)+as.factor(x4))
```