# Example of Convex Optimization Problem

## Regularization terms in Empirical Risk minimization

(Konstantin Burlachenko, burlachenkok@gmail.com )

### About regularization

Regularization solves multi objective optimization problems in some cone and allow trace optimal trade-off curve in math. One possible note about what is regularization is here: https://sites.google.com/site/burlachenkok/articles/what-is-regularization

### About regularization path

In supervised machine learning goal to find from **Class of functions** $F \subseteq \{f: f\ is\ X \to Y\}$ function which minimize **Score Function**. Where **Score function** judge quality of fitted model or quality of selected function. Gold standard is prediction risk (also known as lack of accuracy) which defined as

$$R(F) = E_{xy} L(y, F(x))$$

In fact it's functional, i.e. unknown variable is function *F.* This setup has a lot of math and engineering problems. For example we really don't know joint distribution of (X,Y).

Due to that we don't know joint distribution we create finite sum approximation

$$R(F) = \frac{1}{N} \sum_{i}^{N} L(y_i, F(x_i))$$

Each technics from AI/ML cover some subset of possible functions. And more or less general technics allow to perform *"blowing"* up size of the function space with your hypothesis/approximation/predictors.

Regularization allow make even more. It allow via changing ***parameter*** of regularization trace path indexed by regularization parameter and measure ***how good we are*** via test set, which is our proxy to population. Scientist who research regularization path are J.Friedman, T.Hastie, R.Tibshirani.

### Several Convex regularization functions

| Function | Function name |
|---|---|
| $r(x) = \lambda \lvert x \rvert_2^2$ | Ridge Regression |
| $r(x) = \lambda \lvert x \rvert_1$ | Lasso Regression |
| $r_\beta(x) = \lambda(\sum(\beta - 1)x_j^2 + (2 - \beta)\lvert x_j \rvert), \beta \epsilon [1, 2]$ | Elastic Net(Zou and Hastie, 2005).Bridges lasso -> ridge |
| $r_\gamma(x) = \lambda \lvert x_i \rvert^\gamma$ $\gamma \geq 1$ – penalty is convex | Convex Power Family |