



## リッジ回帰

正則化項として **L2 ノルム** を使ってモデルを訓練する場合、そのモデルを **リッジ回帰 (ridge regression)** と呼ぶ

「ridge」は山の尾根などを表す英単語であり、誤差関数の形状から名付けられている

ref: なっとく！機械学習 p99

ref: 線形代数の半歩先 p139~140

リッジ回帰では、「パラメータのノルムは小さいほどよい」という条件を課すことで、モデルの複雑さを抑える

具体的には、ノルムの二乗 (**L2 ノルム**) すなわち、自分自身との内積

$$\boldsymbol{w}^{\top} \boldsymbol{w}$$

を正則化項として加える

すると、誤差関数は次のように修正される

$$J^{\text{Ridge}}(\boldsymbol{w}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w})^{\top}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}) + \lambda \boldsymbol{w}^{\top} \boldsymbol{w}$$