岭回归与Lasso回归的出现是为了解决线性回归出现的**过拟合**以及在通过正规方程方法求解θ的过程中出现的x转置乘以x不可逆这两类问题的，这两种回归均通过在损失函数中**引入正则化项**来达到目的，具体三者的损失函数对比见下图：

线性回归的损失函数：$J(\theta)=\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)^2$

岭回归的损失函数：$J(\theta)=\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)^2+\lambda\sum_{j=1}^{n}\theta_j^2$

Lasso回归的损失函数：$J(\theta)=\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)^2+\lambda\sum_{j=1}^{n}\left|\theta_j\right|$

其中λ称为正则化参数，如果λ选取过大，会把所有参数θ均最小化，造成欠拟合，如果λ选取过小，会导致对过拟合问题解决不当，因此λ的选取是一个技术活。

岭回归与Lasso回归最大的区别在于岭回归引入的是L2范数惩罚项，Lasso回归引入的是L1范数惩罚项，Lasso回归能够使得损失函数中的许多θ均变成0，这点要优于岭回归，因为岭回归是要所有的θ均存在的，这样计算量Lasso回归将远远小于岭回归。

从概念上讲，我们可以说，Lasso回归（L1）同时做变量选择和参数收缩，而ridge回归只做参数收缩，并最终在模型中包含所有的系数。在有相关变量时，ridge回归可能是首选。此外，ridge回归在用最小二乘估计有更高的偏差的情况下效果最好。因此，选择合适的模型取决于我们的模型的目标。