1 Coordinate Descent Algorithm for LASSO

Consider the following objective function

$$Q(x) = \frac{1}{2}x^T Sx - \alpha^T x + \lambda \sum_{k=1}^{p} |x_k|$$

where S is a $p \times p$ matrix, α is a $p \times 1$ vector, and $\lambda \geq 0$ is a scalar. If we only concentrate on one coordinate, say x_1, then

$$Q(x_1) = \frac{1}{2}s_{11}x_1^2 - (a_1 - x_{(2)}^T S_{12})x_1 + \lambda|x_1| + \lambda \sum_{k=2}^{p} |x_k| + \frac{1}{2}x_{(2)}^T S_{22}x_{(2)} - a_2^T x_{(2)}$$

where $x_{(2)} = (x_2, \ldots, x_p)^T$. Its minimizer is

$$x_1 = \mathrm{sign}(a)\frac{(|a| - \lambda)_+}{s} = \begin{cases} (a - \lambda)/s, & a > \lambda \\ 0, & -\lambda \leq a \leq \lambda \\ (a + \lambda)/s, & a < -\lambda \end{cases}$$

where $s = s_{11}$ and $a = a_1 - x_{(2)}^T S_{12}$.

We can find a minimizer of $Q(x)$ using the coordinate descent algorithm.

- set $x = 0$.
- update the value of x for each coordinate sequentially.
- check convergence. stop until no change in x.

Notice that S is a positive definite matrix, so s_{11} is positive.

Consider the following objective function

$$Q(x) = \frac{1}{2}x^T Sx - \alpha^T x + \sum_{k=1}^{p} \lambda_k |x_k|$$

Denote $W = \text{diag}\{\lambda_k, \ldots, \lambda_k\}$, and let

$$\tilde{S} = W^{-1}SW^{-1}, \quad \tilde{\alpha} = W^{-1}\alpha, \quad \tilde{x} = Wx$$

Therefore, the objective function becomes

$$Q(\tilde{x}) = \frac{1}{2}\tilde{x}^T \tilde{S}\tilde{x} - \tilde{\alpha}^T \tilde{x} + \lambda \sum_{k=1}^{p} |\tilde{x}_k|$$

where $\lambda = 1$. After the minimizer of $Q(\tilde{x})$ is obtained, the minimizer of the original problem is $x = W^{-1}\tilde{x}$.