Efficiently Solving Linear Ridge Regression

- Numerical computation of matrix inversion of $\mathbf{Z}^\top\mathbf{Z}$ is expensive
- Instead we could use singular value decomposition (SVD) to lower the computation cost:

$$\mathbf{Z} = \mathbf{U}\mathbf{D}\mathbf{V}^\top \quad \text{奇异值分解}$$

where:

- $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p)$ is an $n \times p$ orthogonal matrix
- $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, \geq d_p)$ is a $p \times p$ diagonal matrix consisting of the singular values $d_1 \geq d_2 \geq \cdots d_p \geq 0$
- $\mathbf{V}^\top = (\mathbf{v}_1^\top, \mathbf{v}_2^\top, \ldots, \mathbf{v}_p^\top)$ is a $p \times p$ matrix orthogonal matrix

- Proof:

$$\hat{\beta}_\lambda^{\text{ridge}} \boxed{\begin{aligned} &= (\mathbf{Z}^\top\mathbf{Z} + \lambda\mathbf{I}_p)^{-1}\mathbf{Z}^\top\mathbf{y} \\ &= \mathbf{V}\,\underset{j}{\text{diag}}\left(\frac{d_j}{d_j^2+\lambda}\right)\mathbf{U}^\top\mathbf{y} \end{aligned}}$$

提示:

$$\begin{aligned} \mathbf{Z}^\top\mathbf{Z} &= (\mathbf{U}\mathbf{D}\mathbf{V}^\top)^\top(\mathbf{U}\mathbf{D}\mathbf{V}^\top) \\ &= \mathbf{V}\mathbf{D}^\top\mathbf{U}^\top\mathbf{U}\mathbf{D}\mathbf{V}^\top \\ &= \mathbf{V}\mathbf{D}^\top\mathbf{D}\mathbf{V}^\top \\ &= \mathbf{V}\mathbf{D}^2\mathbf{V}^\top \end{aligned}$$