Generalized Ridge regression

Consider the regression model

$$Y = X\beta_* + \varepsilon\,,$$

where $X \in \mathbb{R}^{n\times d}$, β_* is an unknown vector in $\mathbb{R}^d$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$. Define the generalized Ridge estimator by:

$$\widehat{\beta} \in \mathrm{Argmin}_{\beta\in\mathbb{R}^d} \left\{(Y - X\beta)^\top W(Y - X\beta) + (\beta - \beta_0)^\top \Delta(\beta - \beta_0)\right\}\,,$$

where $\beta_0 \in \mathbb{R}^d$, $W \in \mathbb{R}^{n\times n}$ is a diagonal matrix with elements in $[0, 1]$, $\Delta \in \mathbb{R}^{d\times d}$ is a definite-positive matrix.

1. Provide the expression of $\widehat{\beta}$ when $\beta_0 = 0$, $W = I_n$ and $\Delta = \lambda I_d$ where $\lambda > 0$.
2. Solve the optimization problem in the general case.
3. Compute $\mathbb{E}[\widehat{\beta}]$ and show that the estimator is unbiased when $\beta_0 = \beta_*$.
4. Compute $\mathbb{V}[\widehat{\beta}]$ and the mean squared error $\mathbb{E}[\|\widehat{\beta} - \beta_*\|_2^2]$ when $\beta_0 = \beta_*$.
5. Assume that $W = I_n$, $\beta_0 = 0$ and $\Delta = V\Lambda V^\top$ where $X = UDV^\top$ is a singular value decomposition of X and Λ is a diagonal matrix with positive diagonal components. Provide an expression of $\widehat{\beta}$ as a function of U, D, V, Λ and Y.