Structural or analytic derivation

Using held-out covariates to define this full isometry does not alter any fitted function or held-out prediction in the declared model. A data-dependent feature selection, truncated PCA, bandwidth choice or nonlinear preprocessing step is a different operation and must be fitted according to the validation protocol. An unpenalized intercept is retained unchanged in the proof.

Comparative cost and preferred method

With F folds and L penalty values, one thin factorization costs $O(pN^2)$ once, instead of repeating that feature-space reduction FL times. Reduced fits still have their own solver cost. A strong dual-Gram implementation can obtain the same reuse without explicitly storing V when only kernel predictions are required.