which is nonincreasing in λ, and strictly decreases whenever a nonzero fitted direction is present. Its limit at zero is the minimum-norm least-squares coefficient norm.

For lasso, the ℓ_1 norm is monotone as regularization weakens, but the Euclidean norm need not be. Here is a standardized-Gram counterexample for objective $\frac{1}{2}\|y - X\beta\|^2 + \lambda\|\beta\|_1$:

$$G = X^\top X = \begin{pmatrix} 1 & .8 & .8 \\ .8 & 1 & .3 \\ .8 & .3 & 1 \end{pmatrix}, \quad c = X^\top y = (3.7, 3, 3)^\top.$$

The leading principal minors are $1, .36, .014 > 0$, so $G \succ 0$, and $G^{-1}\mathbf{1} = (-15, 10, 10)^\top$. On $.1 \le \lambda \le .11$, the positive vector

$$\beta_\lambda = (2 + 15(\lambda - .1),\ 1 - 10(\lambda - .1),\ 1 - 10(\lambda - .1))^\top$$

satisfies $G\beta_\lambda - c + \lambda\mathbf{1} = 0$, hence is the unique lasso minimizer. Its squared norm is 6.2425 at .11 and 6 at .1: weakening the penalty decreases the Euclidean norm. Such a centred unit-column design exists by taking $X = QG^{1/2}$ with $Q^\top Q = I$, $Q^\top \mathbf{1} = 0$, and $y = XG^{-1}c$. Monotonicity of the ℓ_1 norm follows by adding the two optimality inequalities at two penalty values.

For univariate-response PLS with exact arithmetic, the coefficients are the conjugate-gradient iterates for $G\beta = c$, started at zero, in the Krylov spaces generated by c. Restrict to $\operatorname{ran} G$ if necessary. Write $r_k = c - G\beta_k$, $p_0 = r_0$, $a_k = \|r_k\|^2/(p_k^\top G p_k) > 0$,

$$\beta_{k+1} = \beta_k + a_k p_k, \quad r_{k+1} = r_k - a_k G p_k, \quad p_{k+1} = r_{k+1} + b_k p_k, \quad b_k = \|r_{k+1}\|^2/\|r_k\|^2 \ge 0.$$

For completeness these recurrences give orthogonal residuals and G-conjugate directions by induction. The step size makes $r_{k+1} \perp p_k$; previous conjugacy gives orthogonality to all earlier directions. For $j < k$, $Gp_j = (r_j - r_{j+1})/a_j$ gives $r_{k+1}^\top G p_j = 0$. For $j = k$, the contribution $-\|r_{k+1}\|^2/a_k$ is cancelled by $b_k p_k^\top G p_k$. This closes the induction. In particular $r_k \perp \beta_k$, and

$$\beta_k^\top p_k = b_{k-1}\{\beta_{k-1}^\top p_{k-1} + a_{k-1}\|p_{k-1}\|^2\} \ge 0$$