Conventional solution and proof

Let $d = \bar{x}_1 - \bar{x}_{-1}$, let P_0 be the orthogonal projector onto $\ker W$, and write $q = P_0 d$. Spectral decomposition gives

$$(W + \lambda I)^{-1} d = \lambda^{-1} q + W^{\dagger} d + O(\lambda).$$

Therefore the literal finite limit of the source's unscaled $\delta_\lambda(x)$ generally does not exist: $W = 0, d = x = 1$ gives $\delta_\lambda(x) = 1/\lambda$. The valid limits are

$$\lambda \delta_\lambda(x) \to x^\top q, \qquad \frac{(W + \lambda I)^{-1} d}{\|(W + \lambda I)^{-1} d\|} \to \frac{q}{\|q\|} \quad (q \neq 0).$$

To prove the data-piling claim, write W as a positive multiple of the sum of within-class centred outer products. Then $v \in \ker W$ exactly when $v^\top (x_i - \bar{x}_{y_i}) = 0$ for every training point. Thus each class projects to a single value, and the gap for unit $v \in \ker W$ is $|v^\top d| = |v^\top q| \leq \|q\|$. Equality holds at $v = q/\|q\|$, uniquely up to sign. For prescribed class responses ± 1 and a free intercept, coefficients must satisfy $v^\top d = 2$; Cauchy–Schwarz then gives the unique minimum-norm coefficient $2q/\|q\|^2$, with the intercept chosen to centre the two projected values. This is the maximal piling direction of Exercise 18.8 under its separating-data assumptions. If $q = 0$, no positive-gap zero-within-class-variance direction exists, and the asserted separating conclusion fails.

Structural or analytic derivation

The singular resolvent identifies the nullspace projection as the residue of the $1/\lambda$ pole. The residue, not an unscaled divergent score, is the geometrically meaningful limiting discriminant direction. The extremal gap follows from one Hilbert-space Cauchy–Schwarz inequality.

Comparative cost and preferred method