Total Least Squares

- assumes that not only Y_i, but also X_i are noise
- standard interpretation of OLS: find β such that $Y_i \approx X_i \beta$ with smallest error

alternative interpretation: find corrected responses $\tilde{Y}_i$ and β, such that $\tilde{Y}_i = X_i \beta$ exactly

and then minimize corrections: $\tilde{Y}_i = Y_i + R_i$

$\Rightarrow$ optimization: minimize $\sum_i R_i^2 \Rightarrow$ OLS

- if the X_i are also noisy, correct them as well: $\tilde{X}_i = X_i + E_i$, such that $\tilde{Y}_i = \tilde{X}_i \cdot \beta$ exactly for the optimal β

$Y_i + R_i = (X_i + E_i)\beta$ and minimize the corrections R and E

form correction matrix $M = [E\ R]$ concatenation of E and R

TLS objective: $\boxed{\hat{\beta}, \hat{E}, \hat{R} = \arg\min_{\beta, E, R} \|M\|_F^2 \quad \text{s.t.} \quad (X+E)\beta = Y+R}$

$\hat{=}$ constrained optimization problem

- form the data matrix $Z = [X \ Y]$ (concatenate X and Y) and compute its singular value decomposition (SVD): $Z = U \Lambda V^T$

U: orthonormal $N \times (D+1)$; Λ: diagonal ~~$D \times D$~~ $(D+1)\times(D+1)$; V: orthon. ~~$D \times D$~~ $(D+1)\times(D+1)$

solution of TLS objective: $\min \|M\|_F^2$ s.t. ... $= \min_{j} \lambda_j^2$

in words: the minimal correction equals the smallest singular vector of Z

[alternative: can also form $Z^T Z$ and find the smallest eigenvalue]

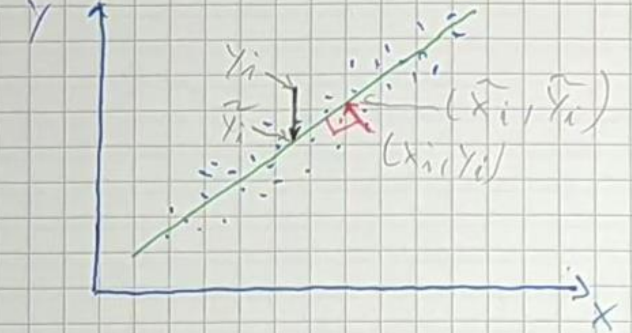
construct β from the SVD solution: let v_j be the singular vector for the smallest singular value λ_j: split v_j into the first D dimension (for the X part) and the last dimension (for the Y part)

$$v_{\hat{j}} = \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} \begin{matrix} \text{] vector size } D \\ \leftarrow \text{scalar} \end{matrix} \quad \text{with } \hat{j} = \arg\min_j \lambda_j$$

$$\boxed{\hat{\beta} = -\frac{\alpha}{\gamma}}$$

(singular vector $\hat{=}$ columns of V $\hat{=}$ eigenvectors of $Z^T Z$)

• example : 2D line fitting



— OLS correction: correct only $\tilde{Y}_i = Y_i + e_i$

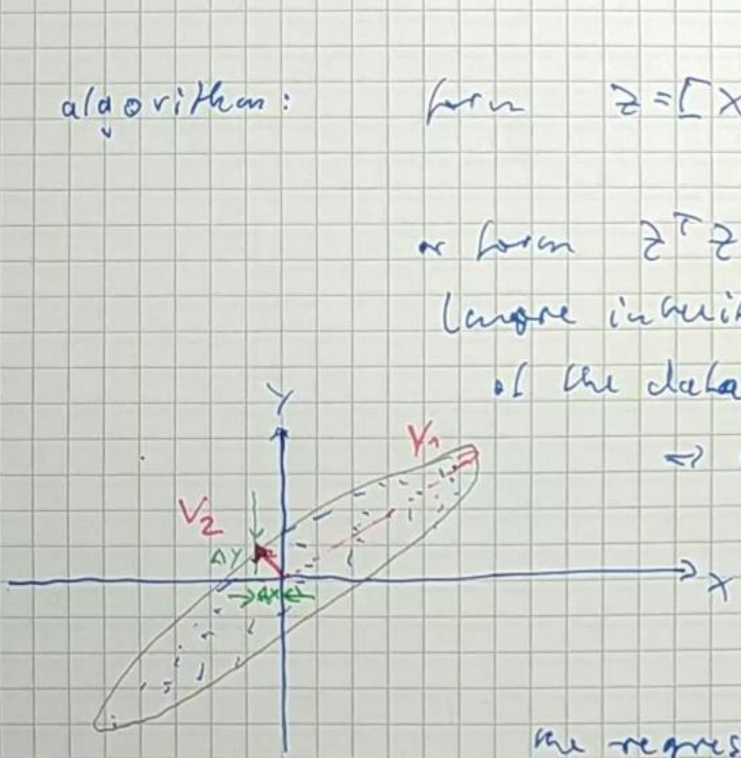
— TLS correction: $\tilde{X}_i = X_i + E_i$, $\tilde{Y}_i = Y_i + e_i$

turns out that in the optimal TLS solution, the correction is always perpendicular to the regression line (in contrast, OLS correction is parallel to the Y axis)

algorithm: form $Z = [X, Y]$, find smallest singular value and corresponding singular vector

or form $Z^T Z$ and find smallest eigenvector

(more intuitive because $S = Z^T Z$ is the scattermatrix of the data, provided that X and Y <u>are centered</u>)



$\Rightarrow$ data are approximated by an ellipse, and smallest eigenvector is smallest axis of the ellipse

$$V_2 = \begin{bmatrix} \alpha \\ \gamma \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix},$$ this is the <u>normal</u> of the regression line (or regression plane in higher dims)

regression line is perpendicular to normal $v_2^{\perp} = \begin{bmatrix} -\Delta Y \\ \Delta X \end{bmatrix}$

$$\Rightarrow \boxed{\beta = -\frac{\Delta X}{\Delta Y}}$$ slope of TLS regression line