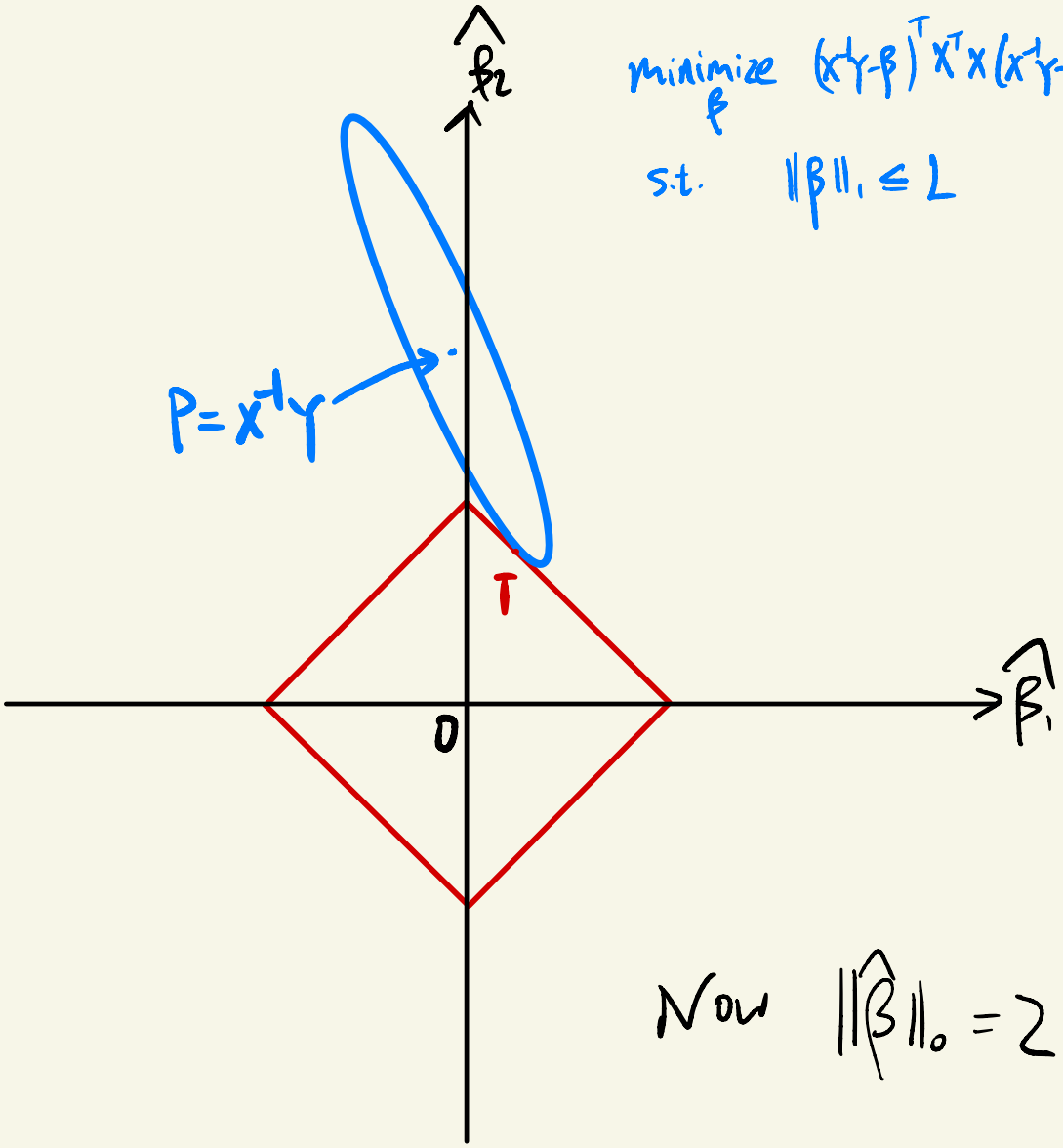
- Example where $\|\hat{\beta}\|_1 \uparrow$ but $\|\hat{\beta}\|_0 \downarrow$

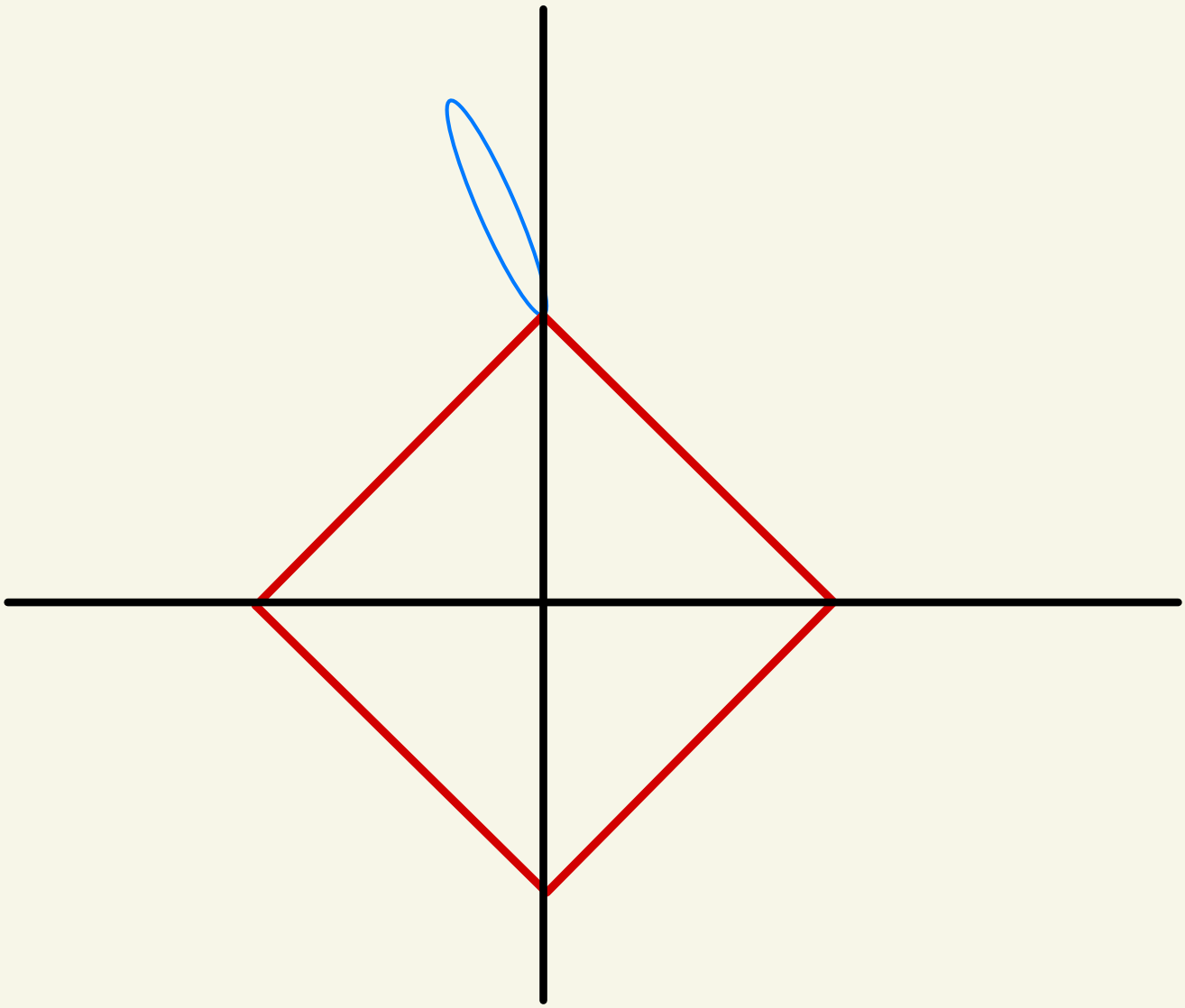
$P := X^{-1}Y$. T is the tangent point. Line PT has a smaller slope (in absolute value) than PO



$$\underset{\beta}{\text{minimize}} \; (X^{-1}Y-\beta)^T X^T X (X^{-1}Y-\beta)$$

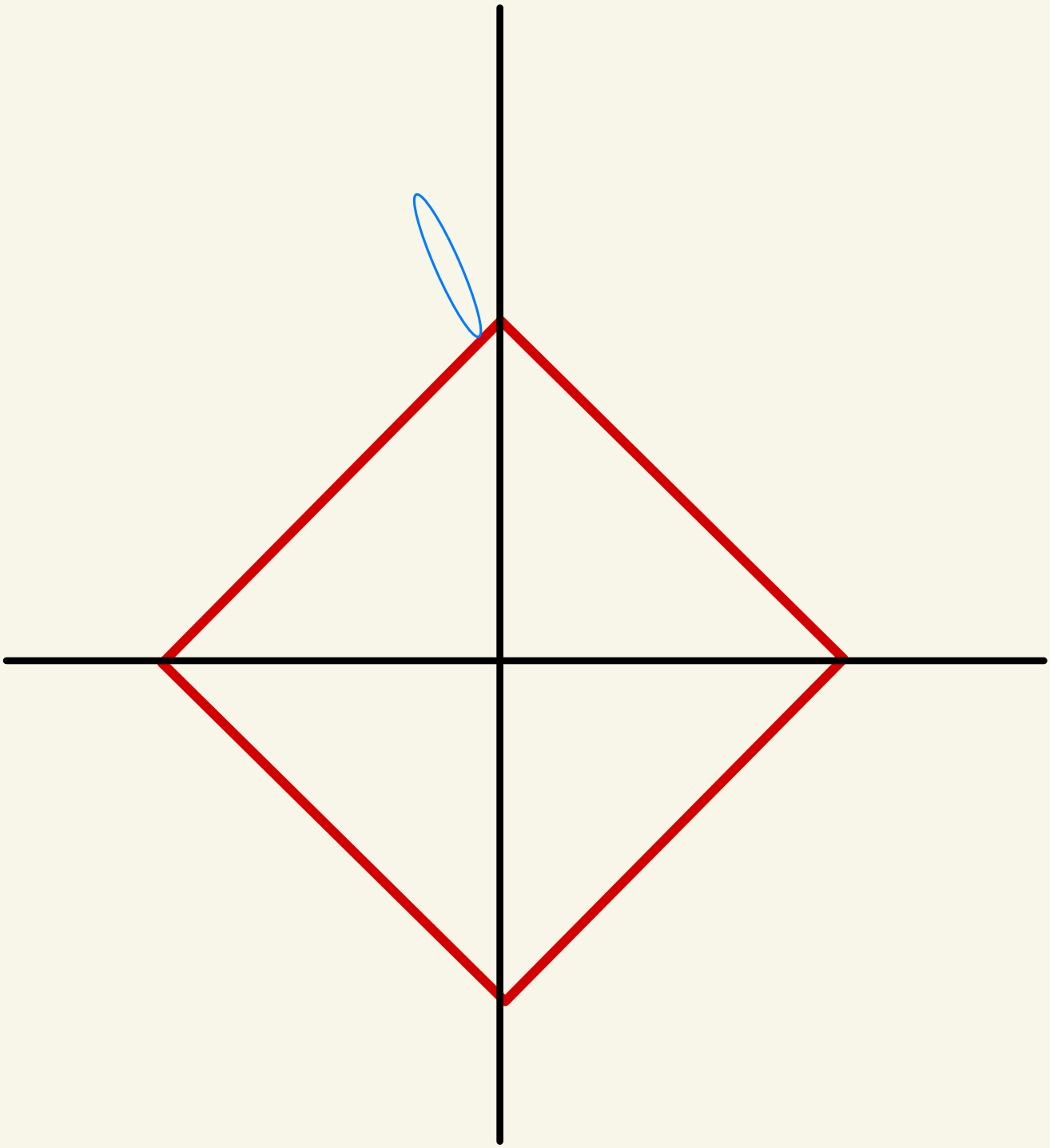
$$\text{s.t.} \quad \|\beta\|_1 \le L$$

Now $\|\hat{\beta}\|_0 = 2$



Increase $\|\hat{\beta}\|_1$.

Now $\|\hat{\beta}\|_0 = 1$



Further increase $\|\hat{\beta}\|_1$

Now $\|\hat{\beta}\|_0 = 2$ again