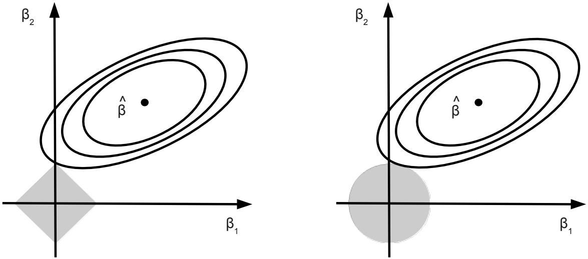
# Recap

Lasso:

$$\hat{\beta}(\lambda) = \text{argmin}_{\beta}(\|Y - X\beta\|_2^2/n + \lambda\|\beta\|_1)$$

- sparse estimator



- convex optimization

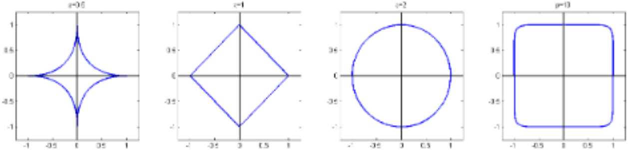


**Figure 1:** Unit circles for several Minkowski-$p$-norms $\|\mathbf{x}\|_p$: from left to right $p = 0.5$, $p = 1$ (Manhatten), $p = 2$ (Euclidean), $p = 10$.

Figure from Lange, Zühlke, Holz, Villmann (2014)

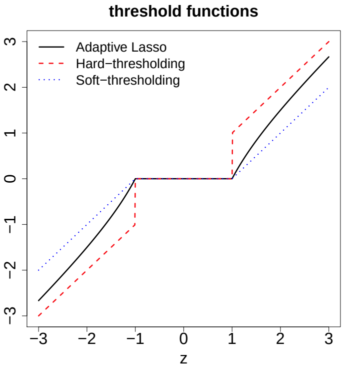
convex: $\ell_p$-norm with $p \geq 1$
sparse: $\ell_p$-norm with $p \leq 1$ (need “edges” in the ball)
$\Longrightarrow$ $p = 1$(*Lasso*) for sparse and convex estimator

# Orthonormal design: explicit solution

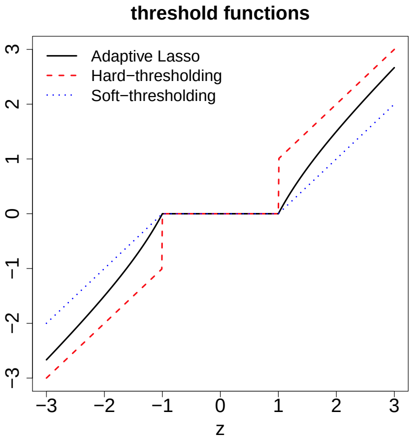
$$X^T X/n = I_{p \times p}$$

Lasso = soft-thresholding of ordinary least squares

$$\hat{\beta}_j(\lambda) = g_\lambda(Z_j), \quad Z_j = (X^T Y)_j/n = \hat{\beta}_{\mathrm{OLS},j},$$
$$g_\lambda(z) = \mathrm{sign}(z)(|z| - \lambda/2)_+$$



# Orthonormal design: explicit solution




$\rightsquigarrow$ Lasso (blue dashed line) exhibits bias!

(Note that OLS is unbiased)