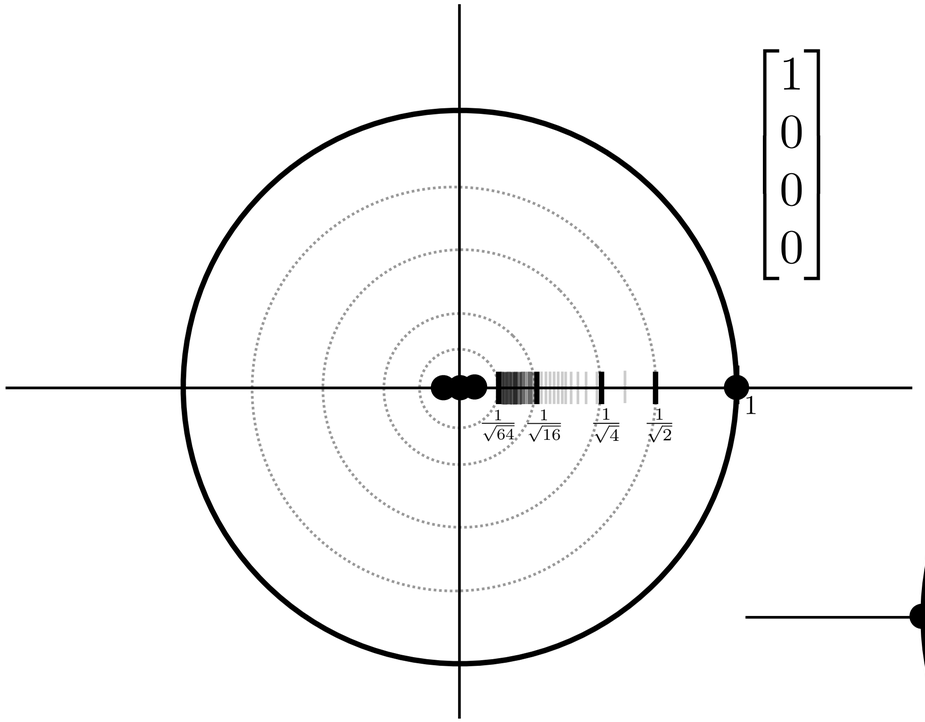
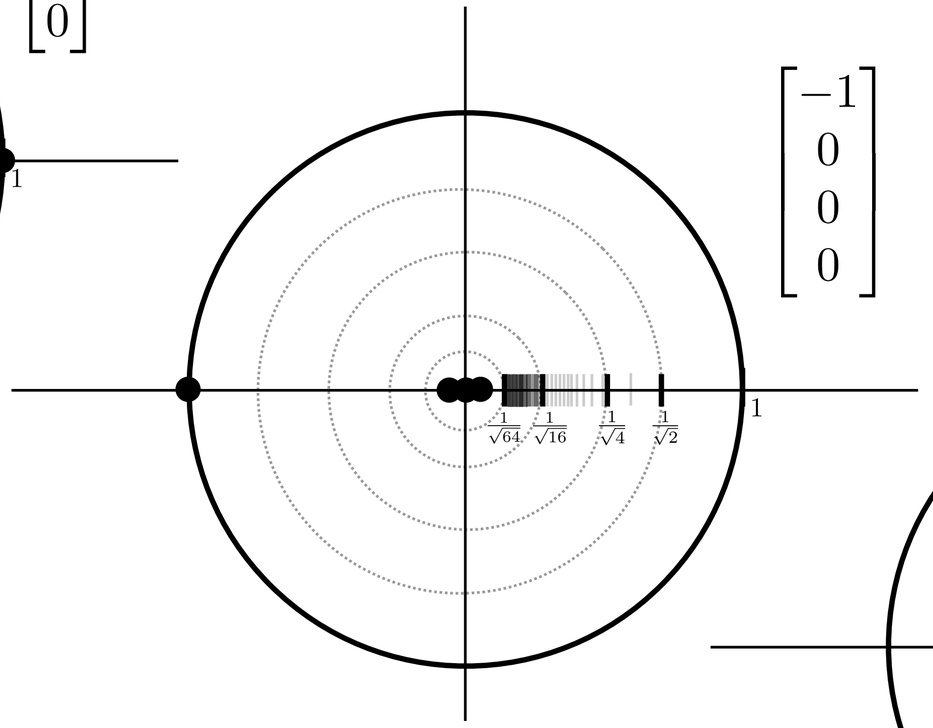
Inner Products in High Dimensions

Dan Calderone - Spring 22

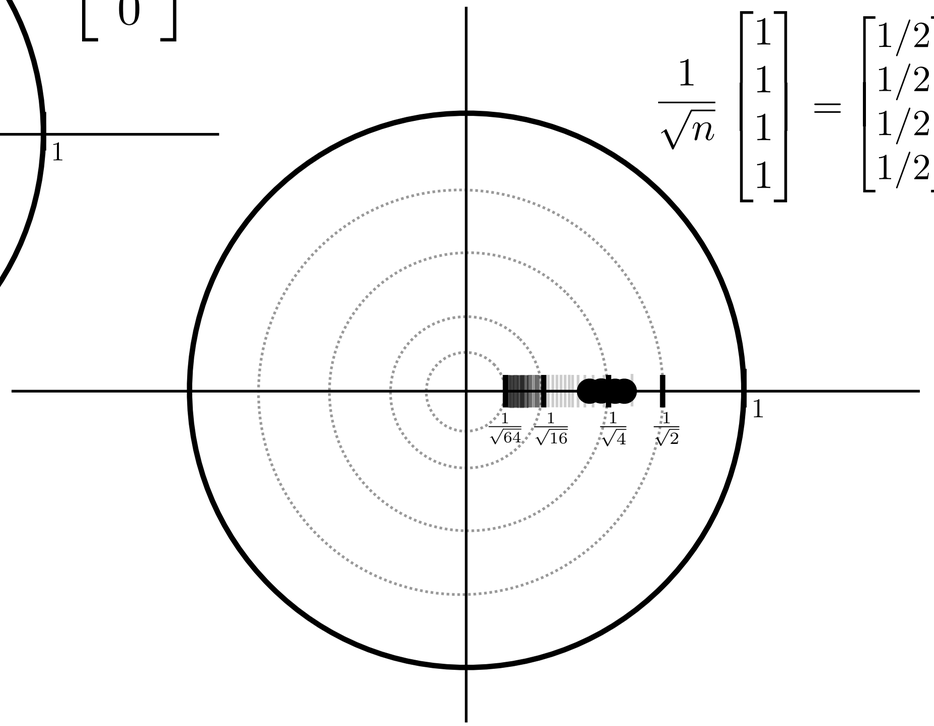
Unit vectors (in high dimensions)



$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$



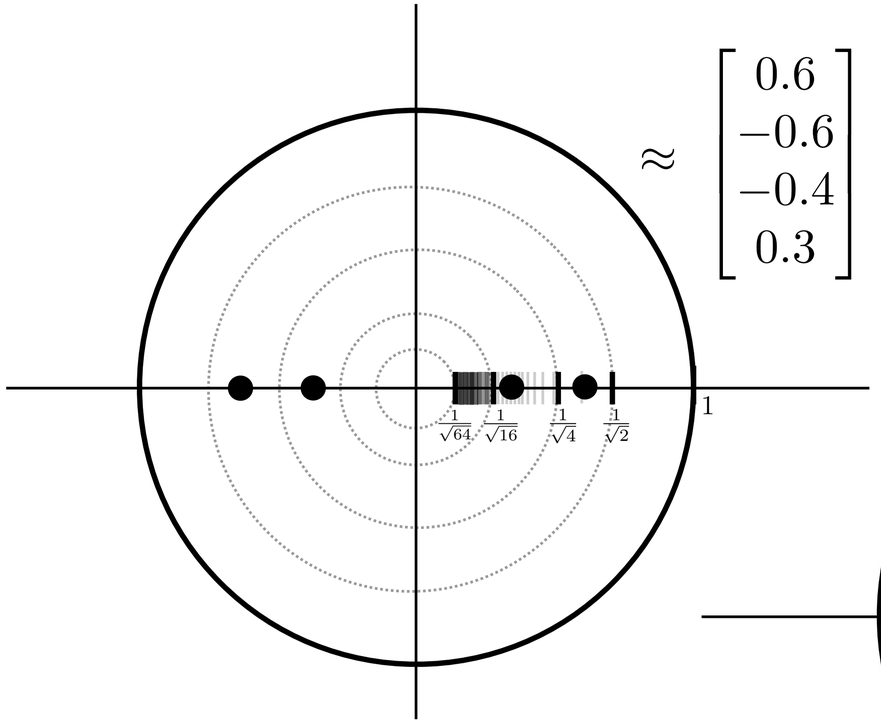
$$\begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$



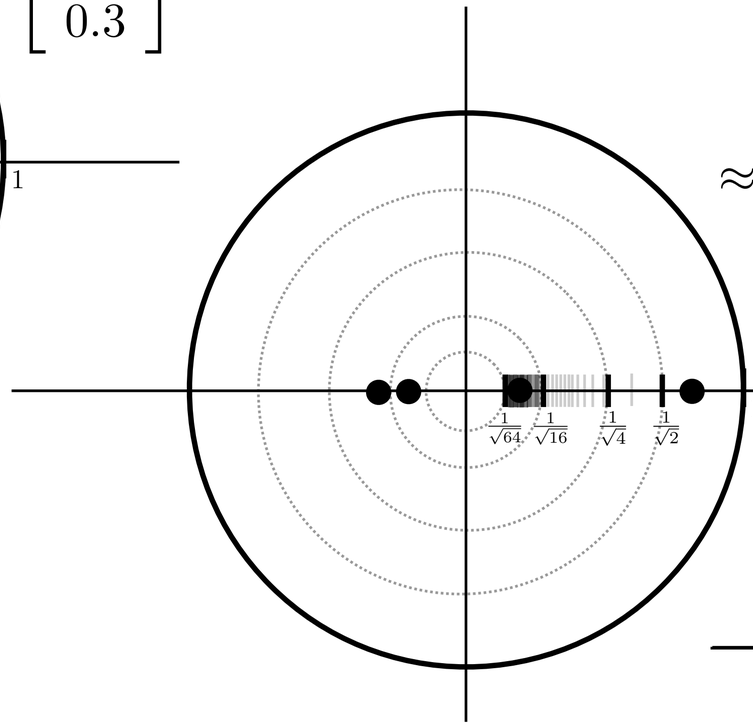
$$\frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

$$n = 4$$

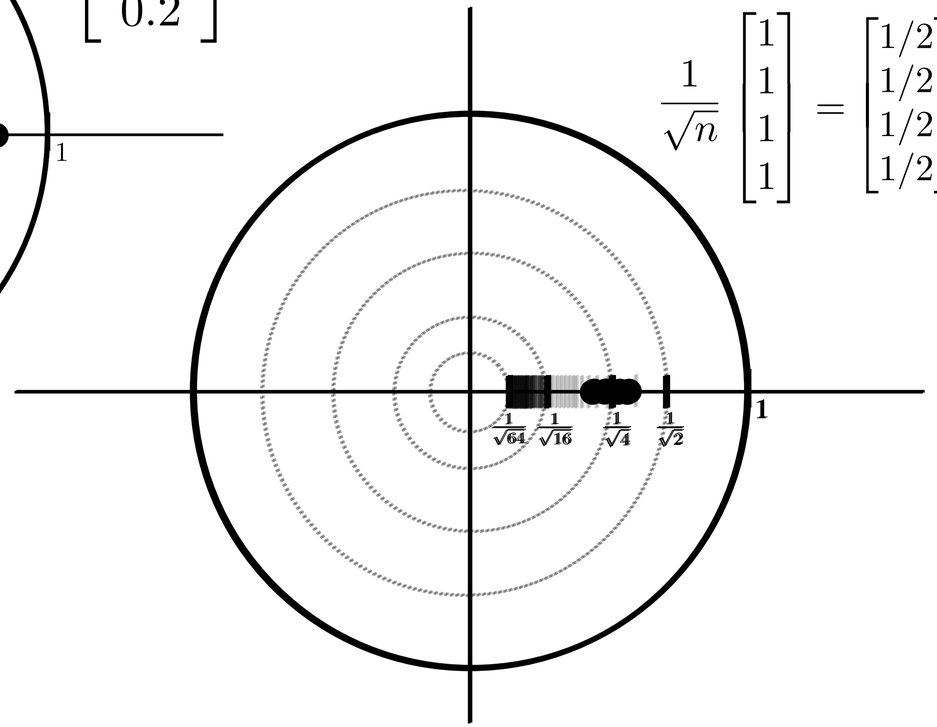
Unit vectors (in high dimensions)



$$\approx \begin{bmatrix} 0.6 \\ -0.6 \\ -0.4 \\ 0.3 \end{bmatrix}$$



$$\approx \begin{bmatrix} 0.9 \\ -0.3 \\ -0.2 \\ 0.2 \end{bmatrix}$$



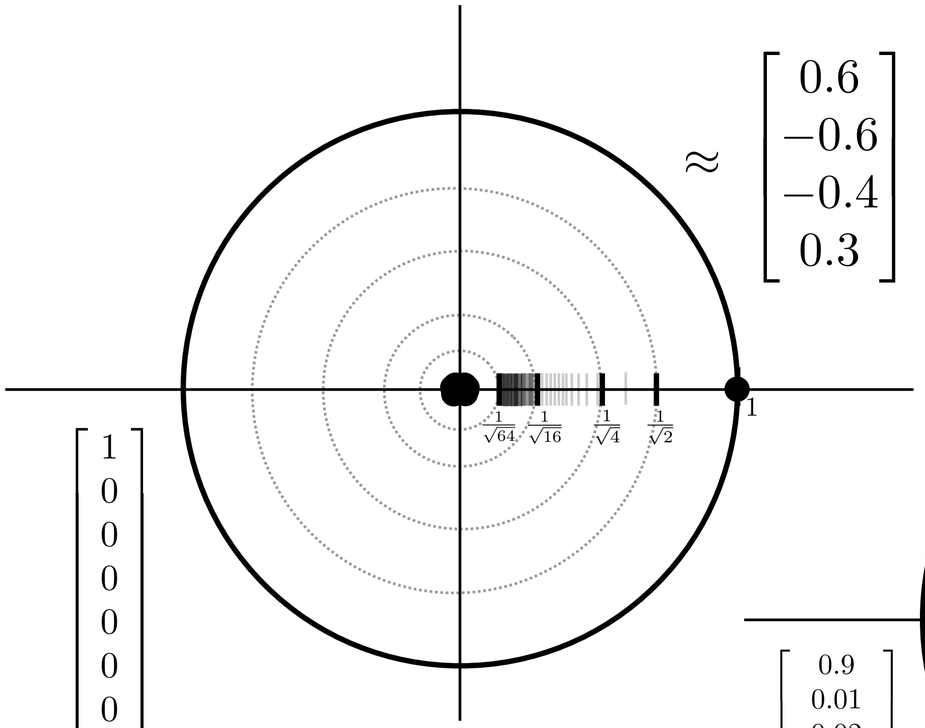
$$\frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$$

$$n = 4$$

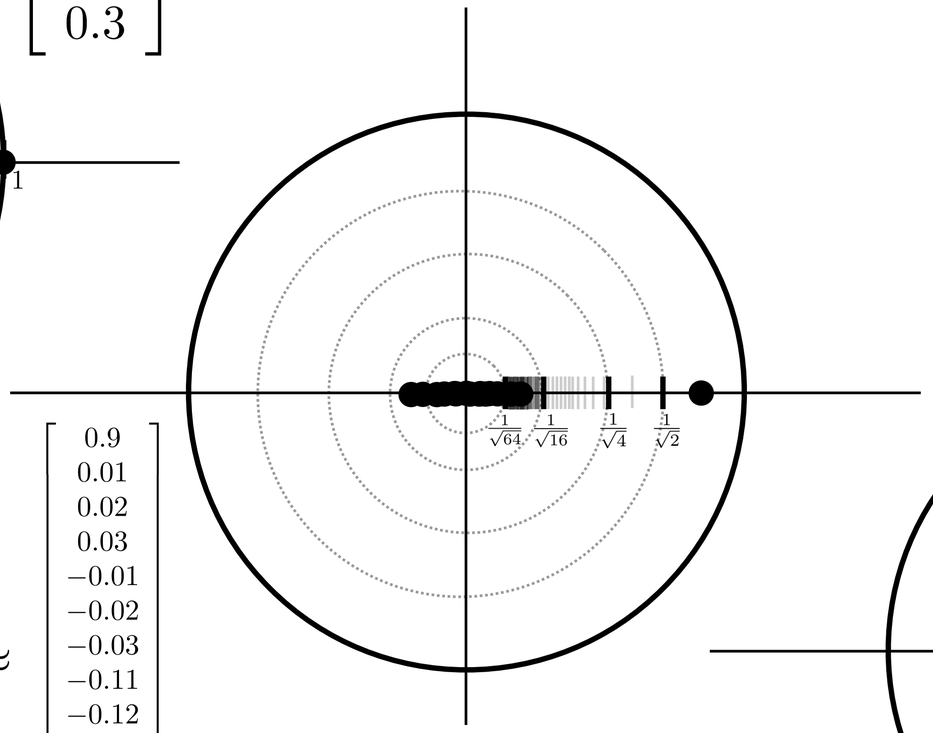
Unit vectors (in high dimensions)

$$n = 16$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

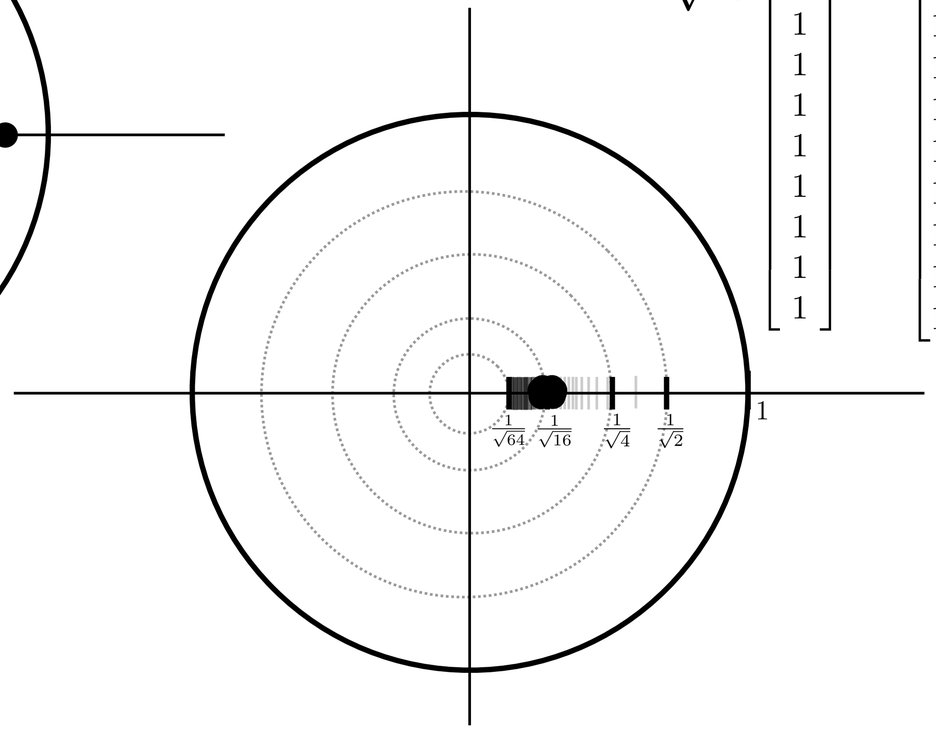


$$\approx \begin{bmatrix} 0.6 \\ -0.6 \\ -0.4 \\ 0.3 \end{bmatrix}$$

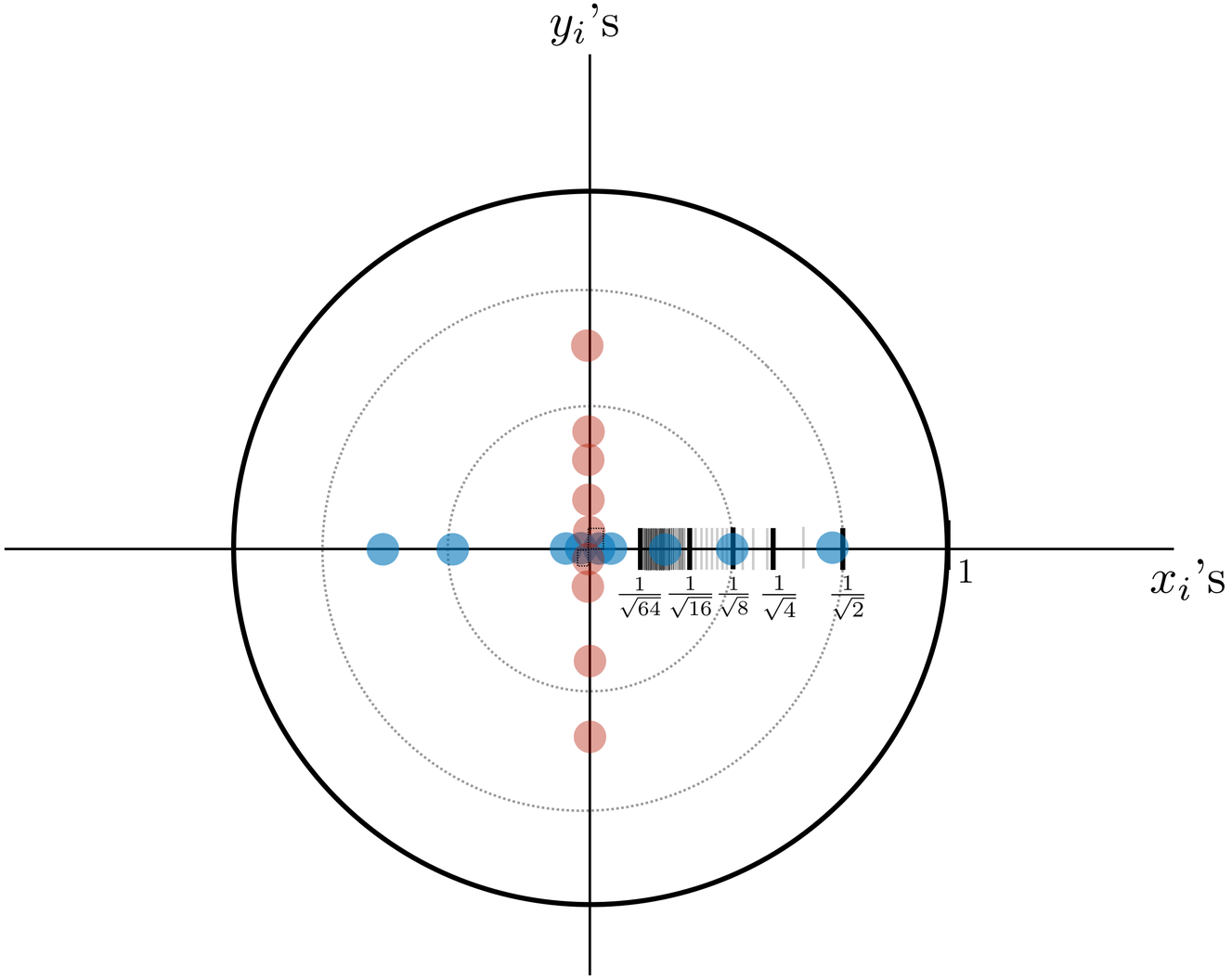


$$\approx \begin{bmatrix} 0.9 \\ 0.01 \\ 0.02 \\ 0.03 \\ -0.01 \\ -0.02 \\ -0.03 \\ -0.11 \\ -0.12 \\ -0.13 \\ 0.11 \\ 0.12 \\ 0.13 \\ 0.14 \\ 0.15 \\ -0.14 \end{bmatrix}$$

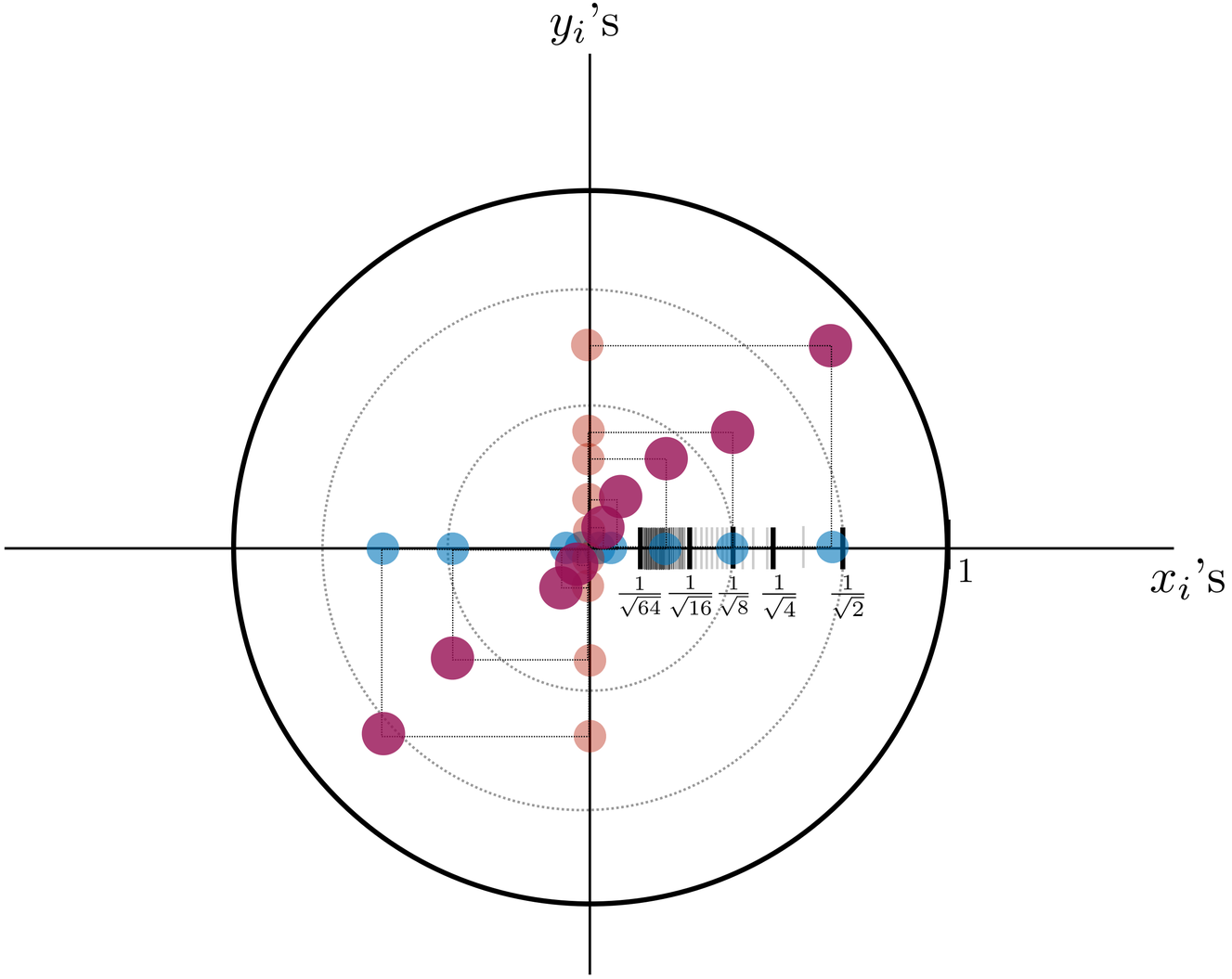
$$\frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$



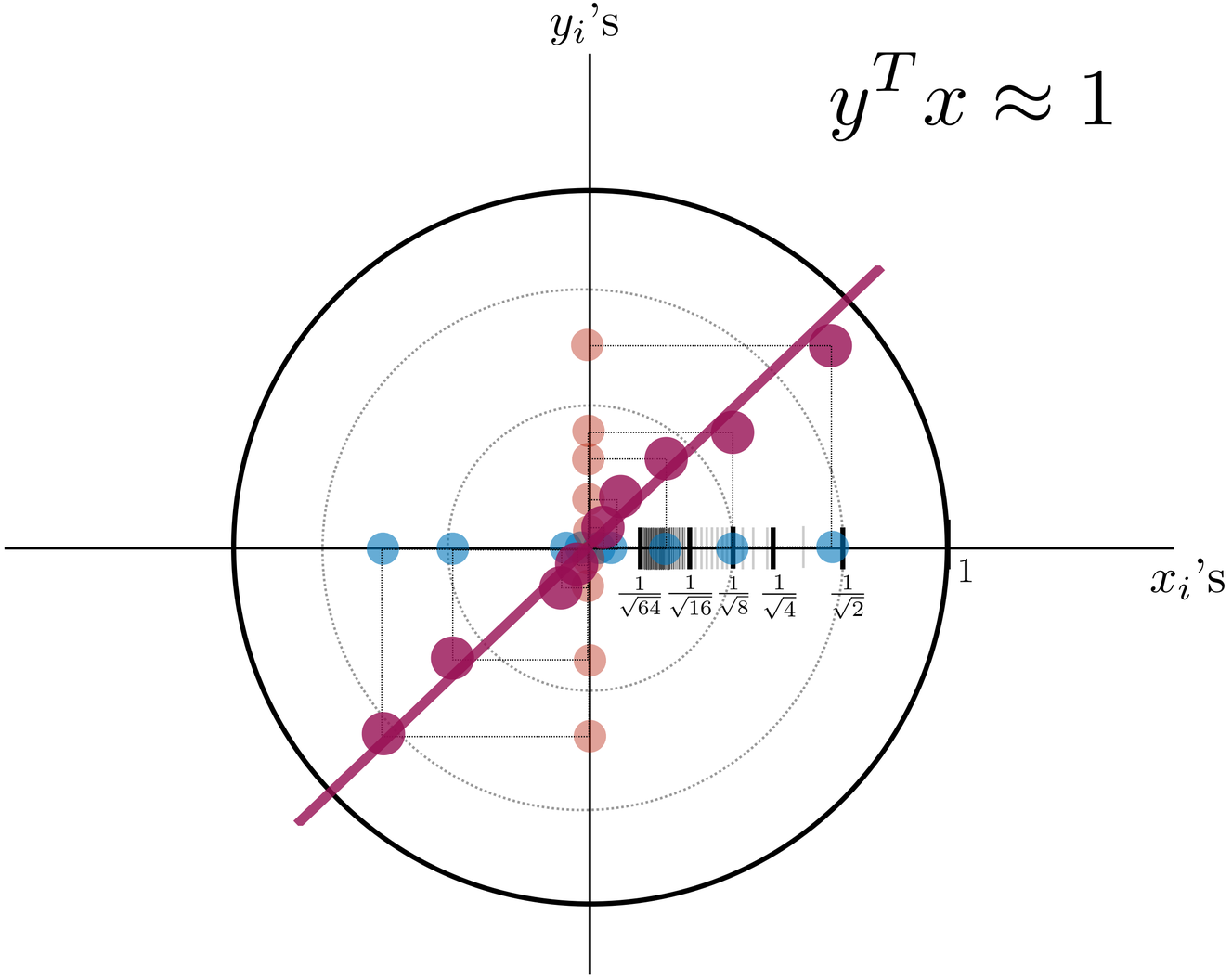
Inner products (in high dimensions)



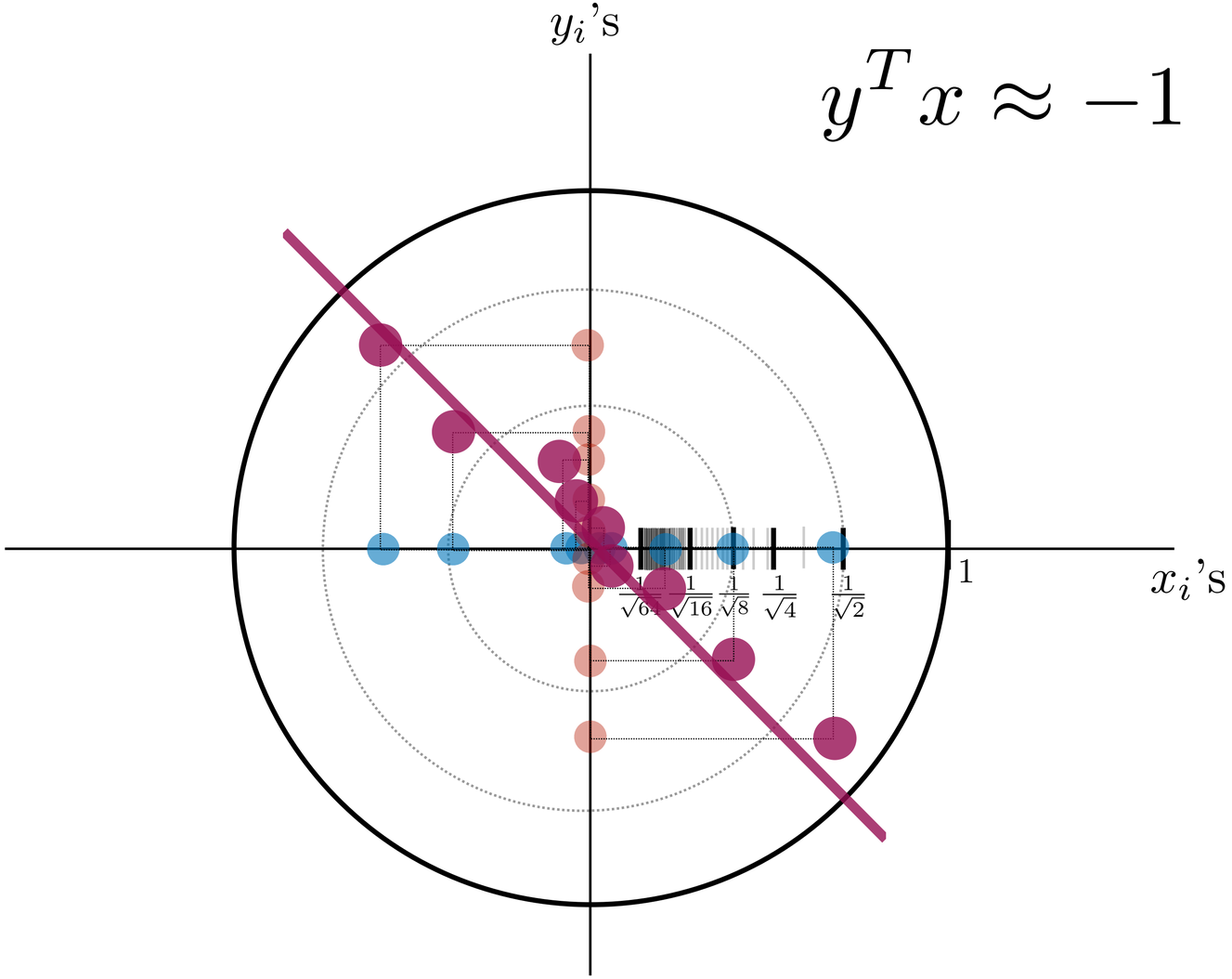
Inner products (in high dimensions)



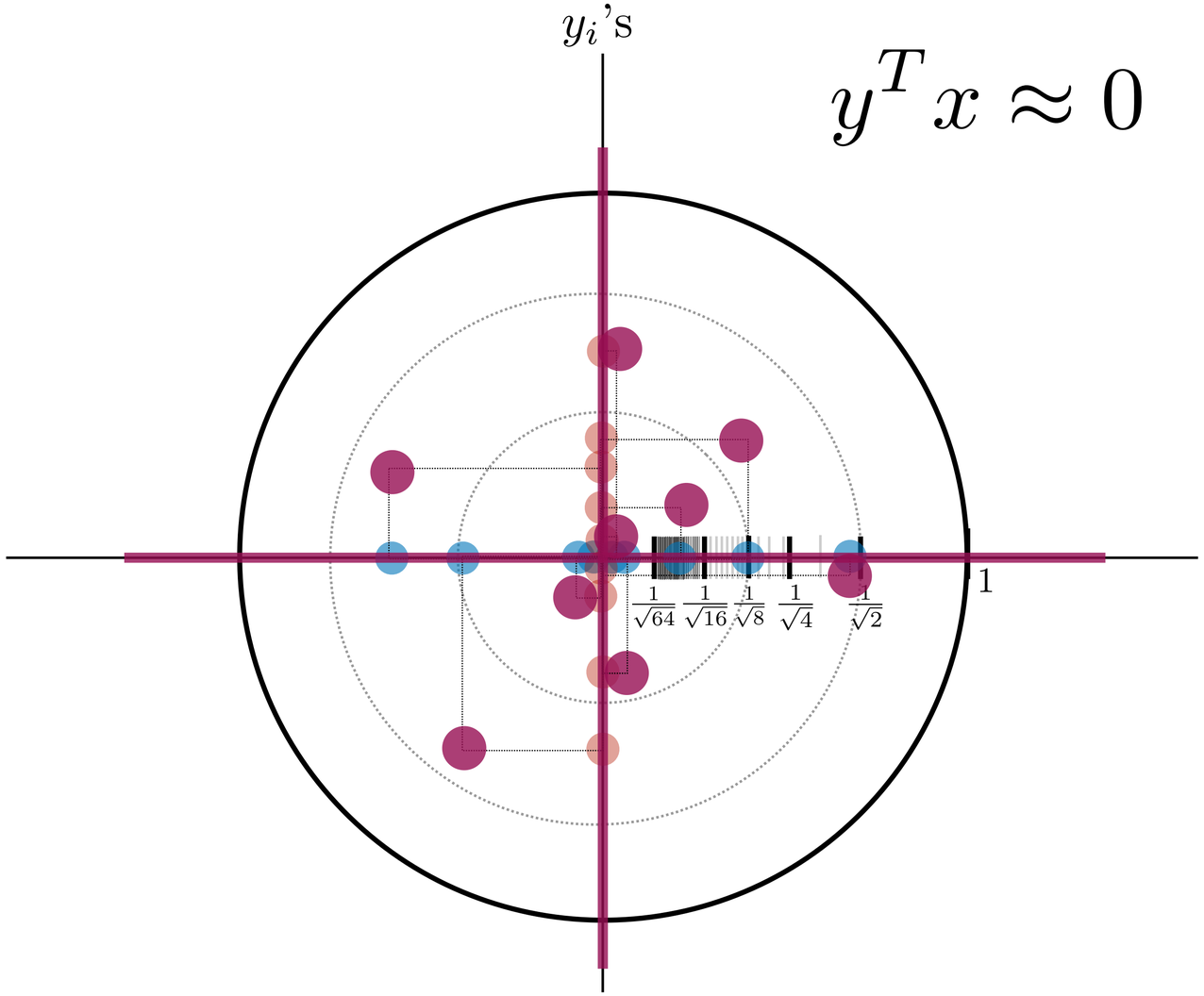
Inner products (in high dimensions)



Inner products (in high dimensions)

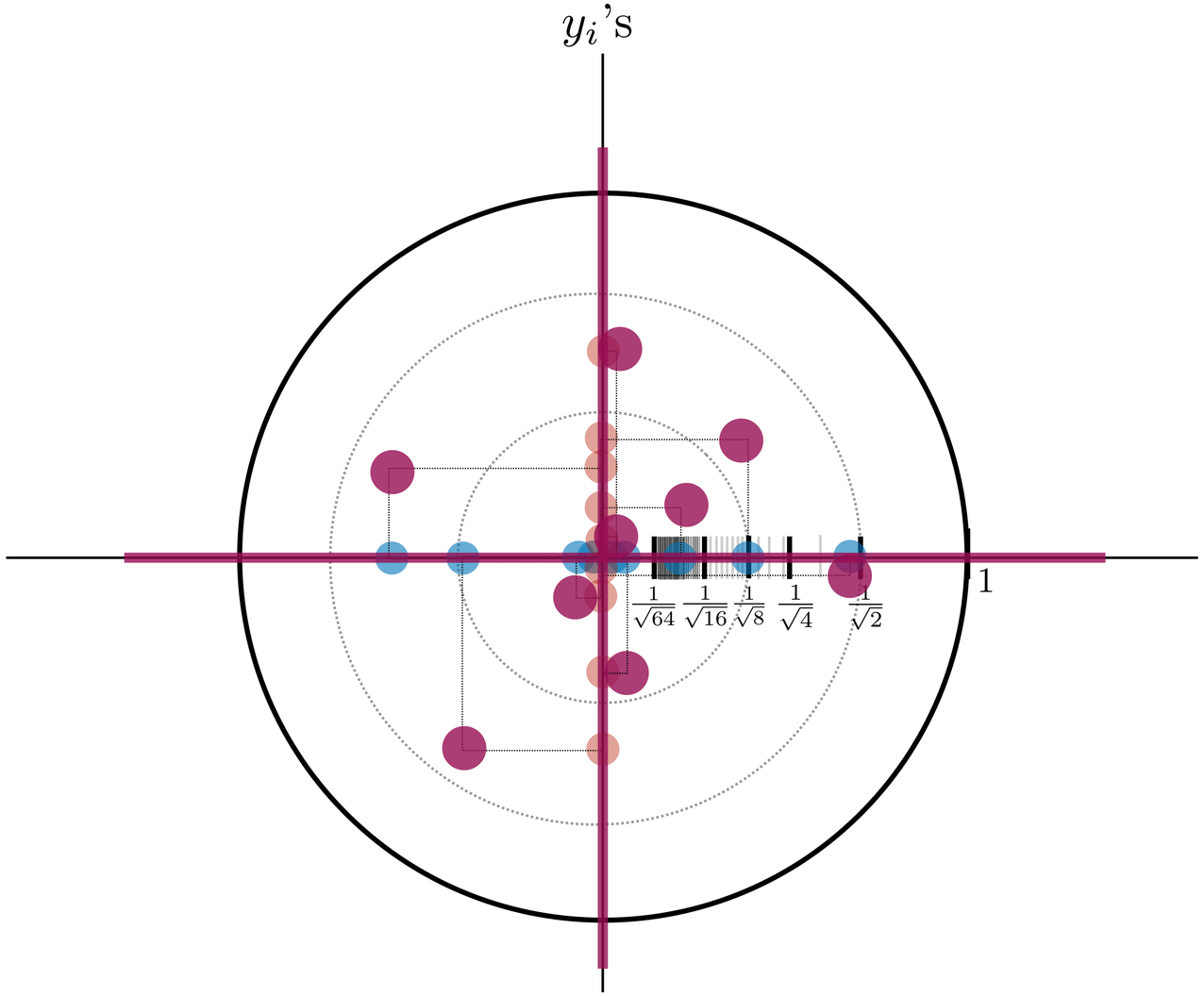


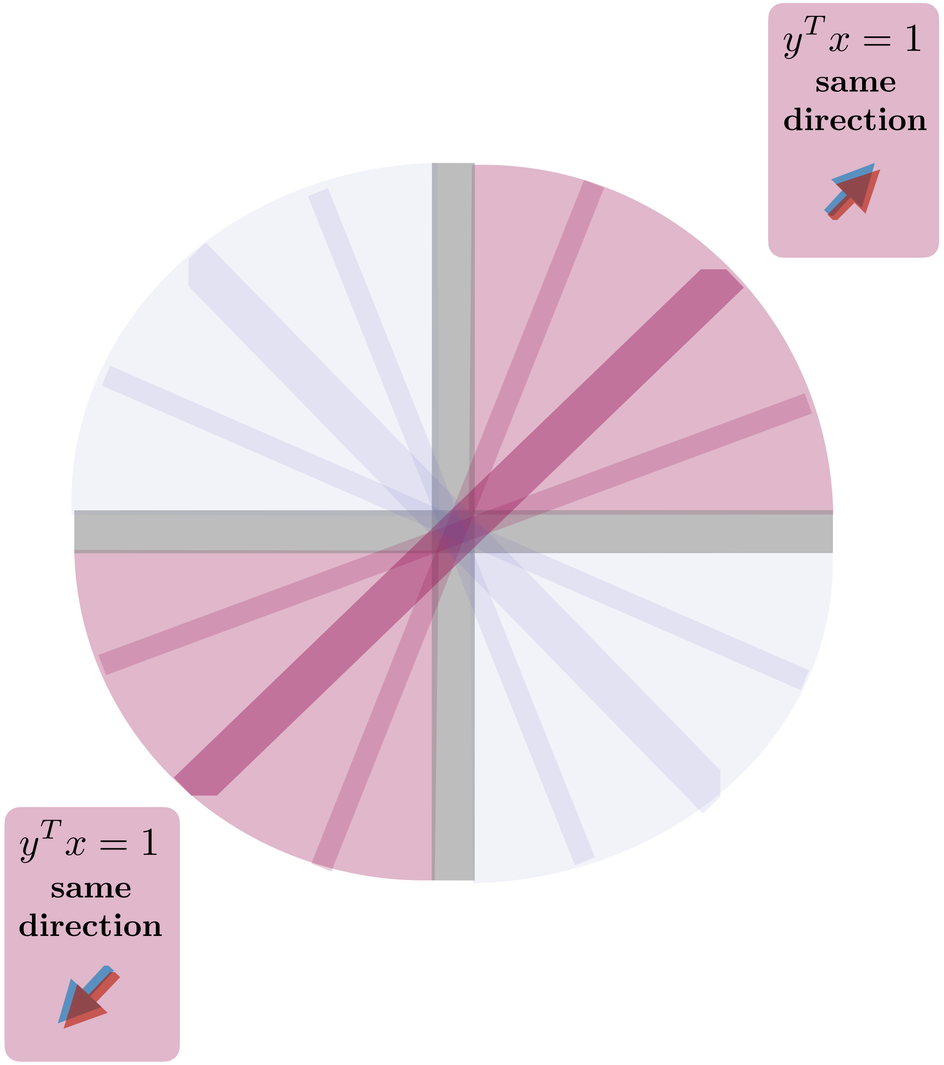
Inner products (in high dimensions)



Inner products (in high dimensions)

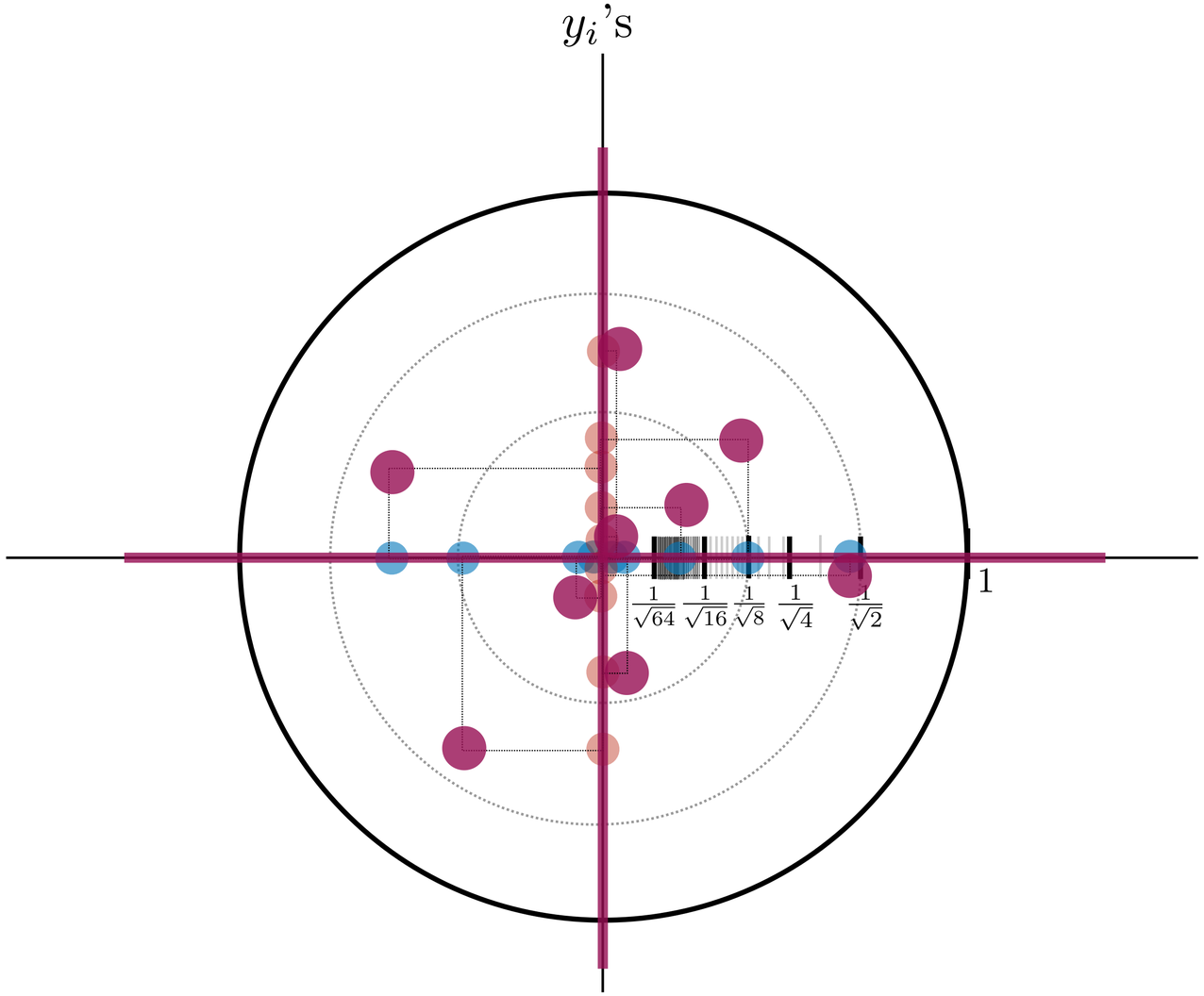
$$y^T x \geq 0$$

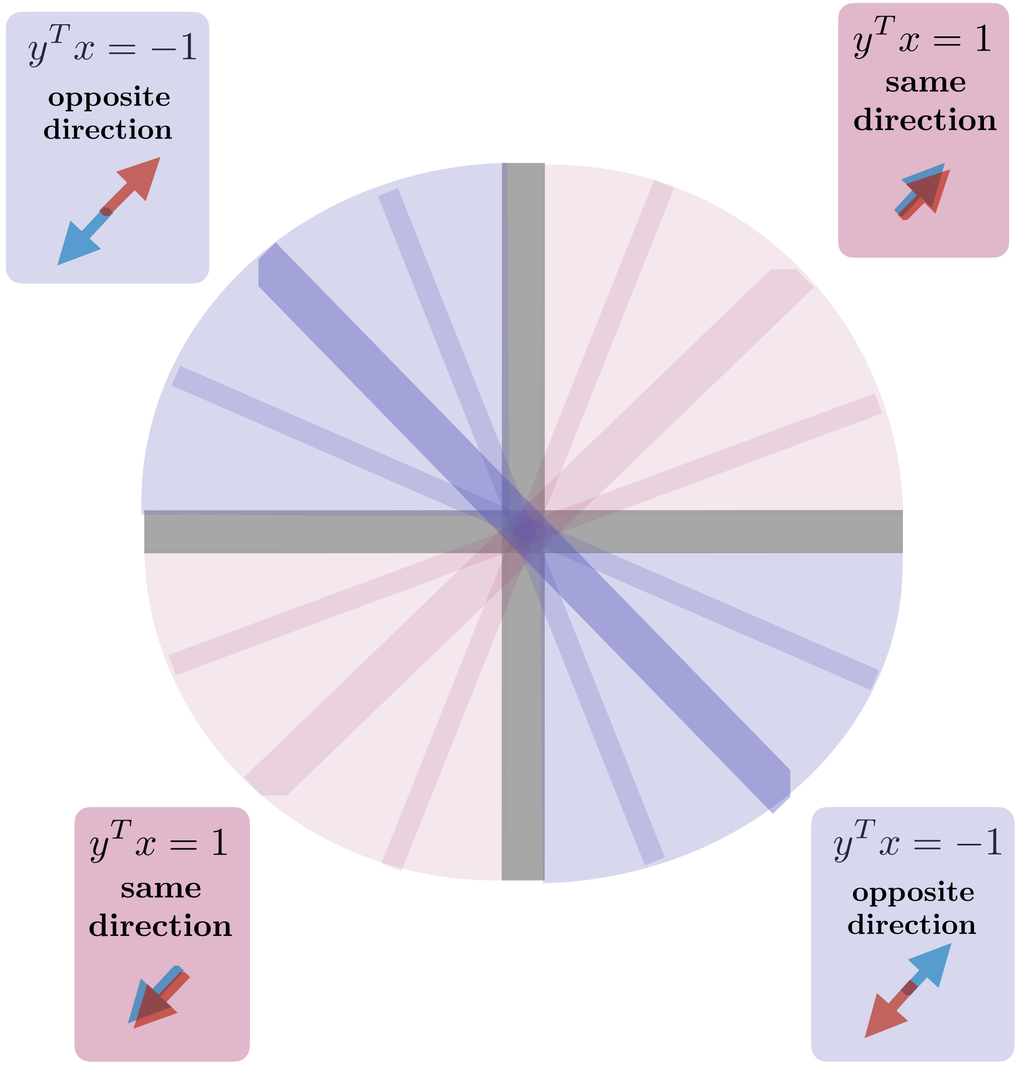




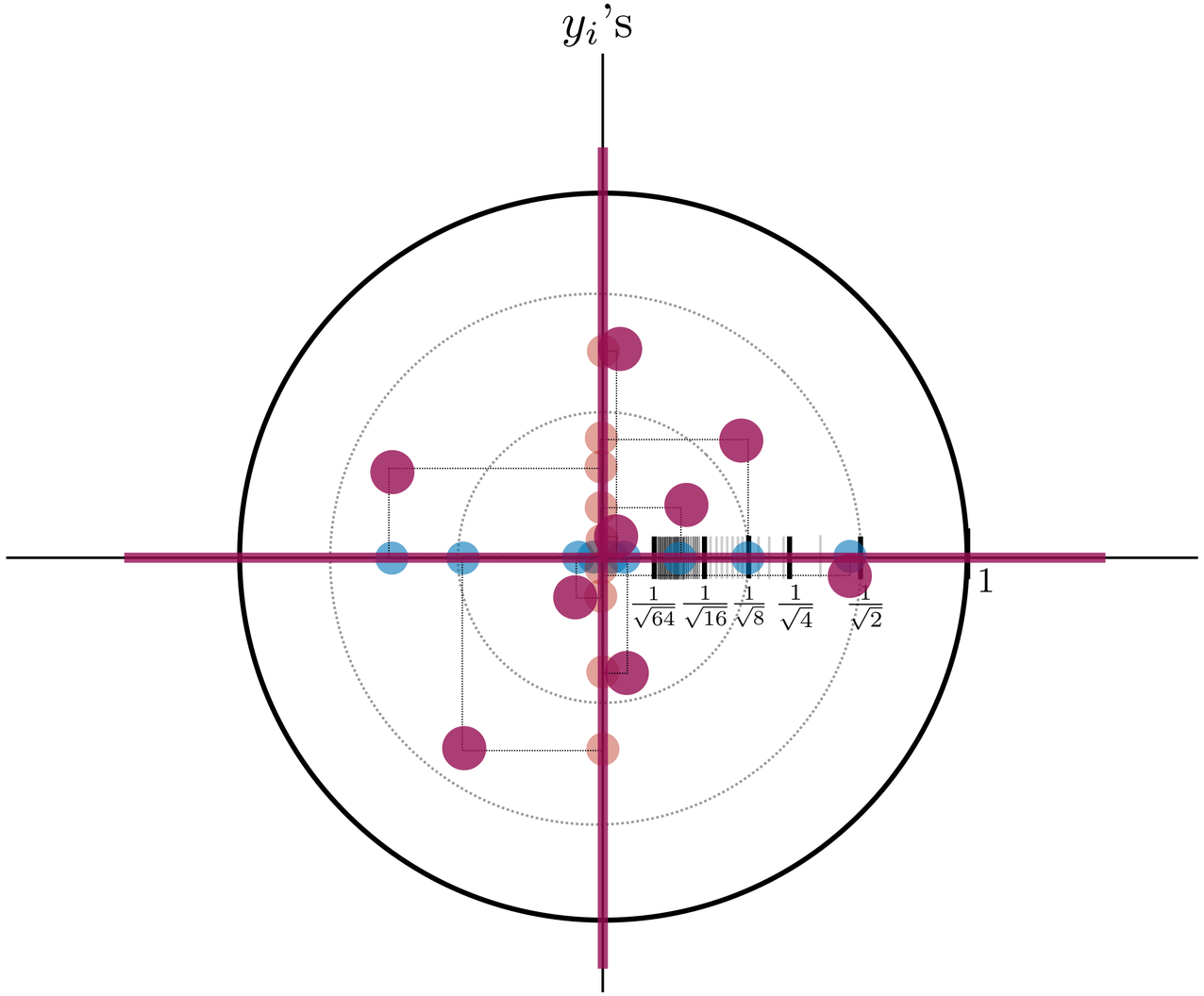
Inner products (in high dimensions)

$$y^T x \leq 0$$

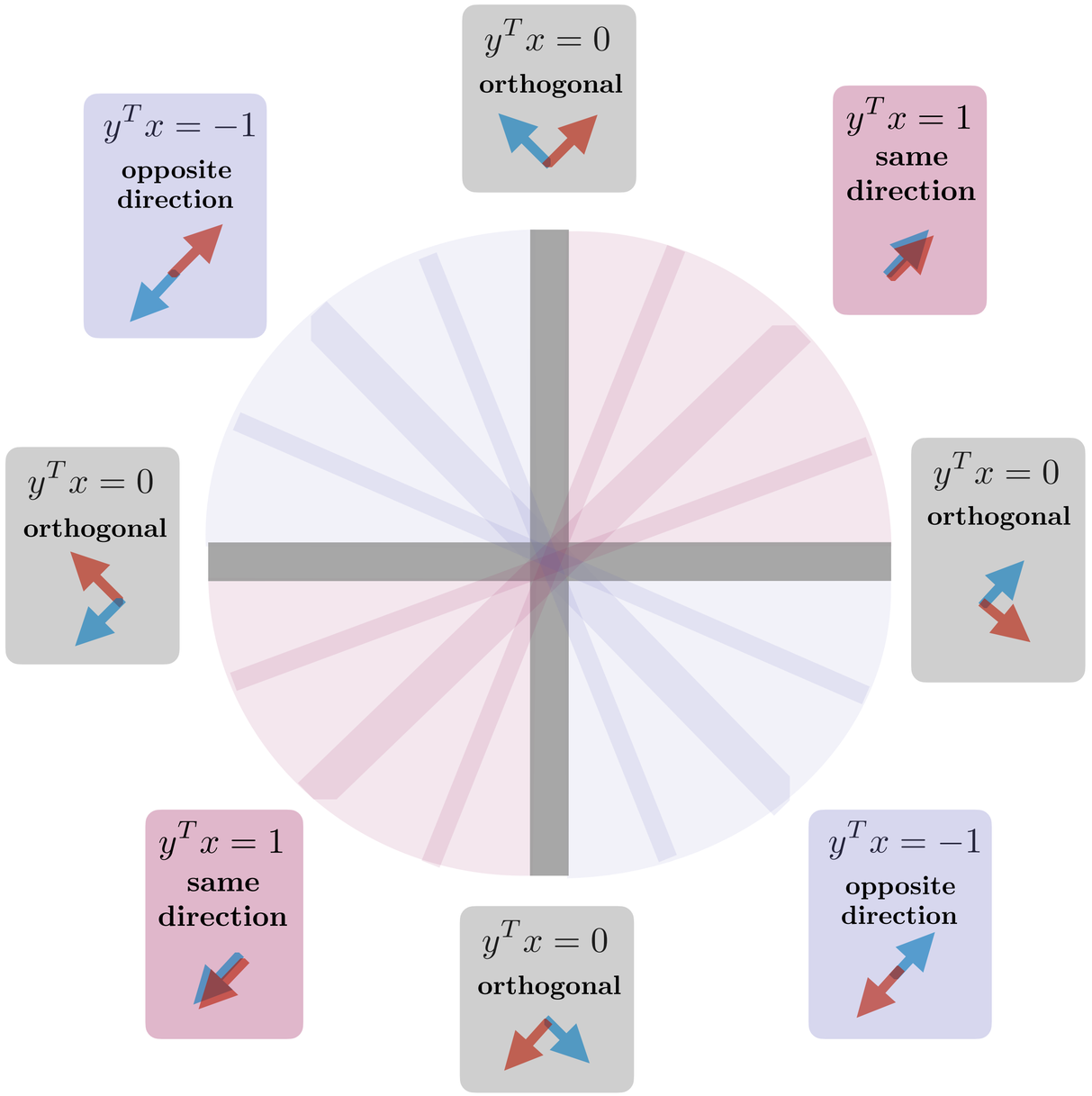




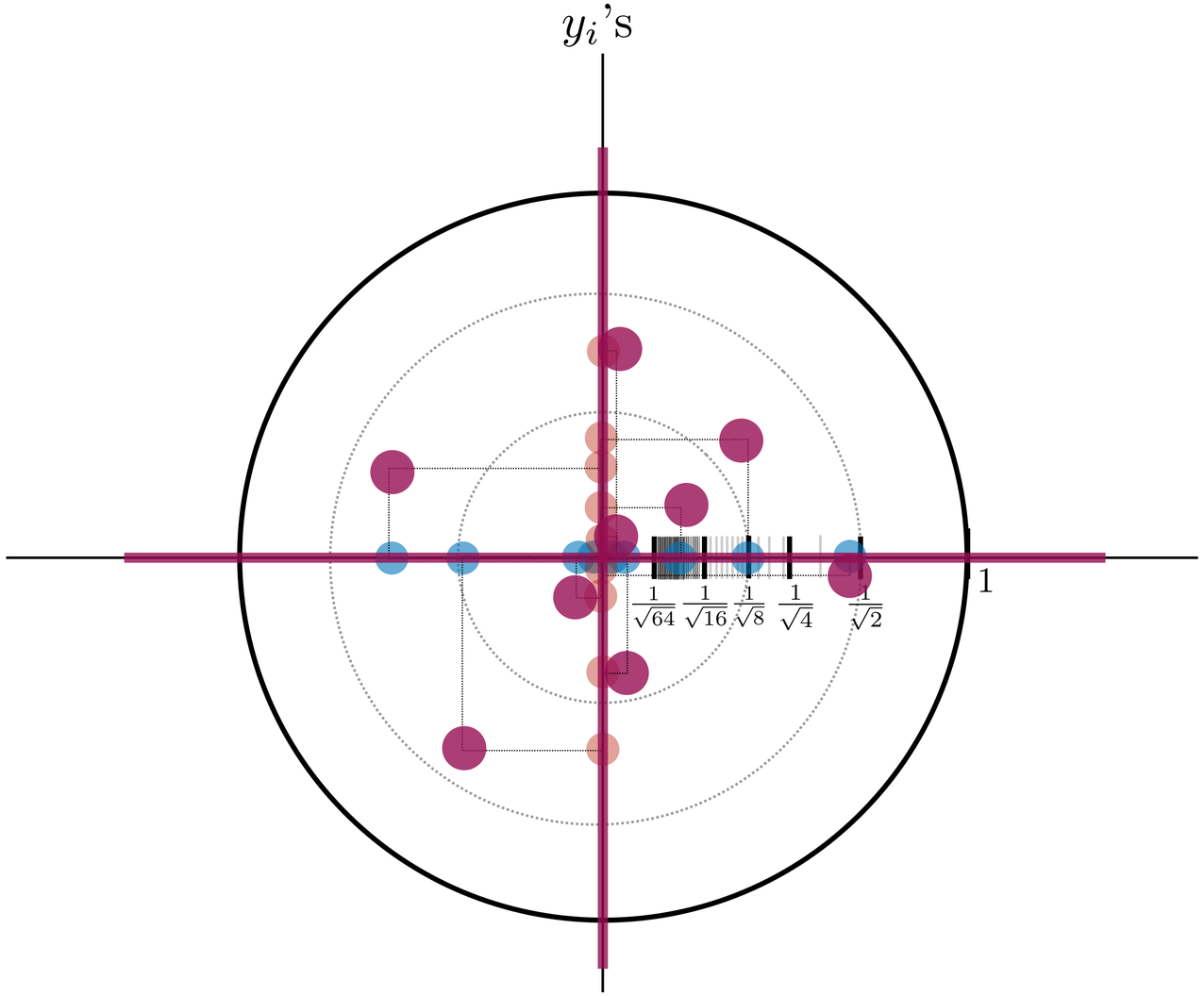
Inner products (in high dimensions)

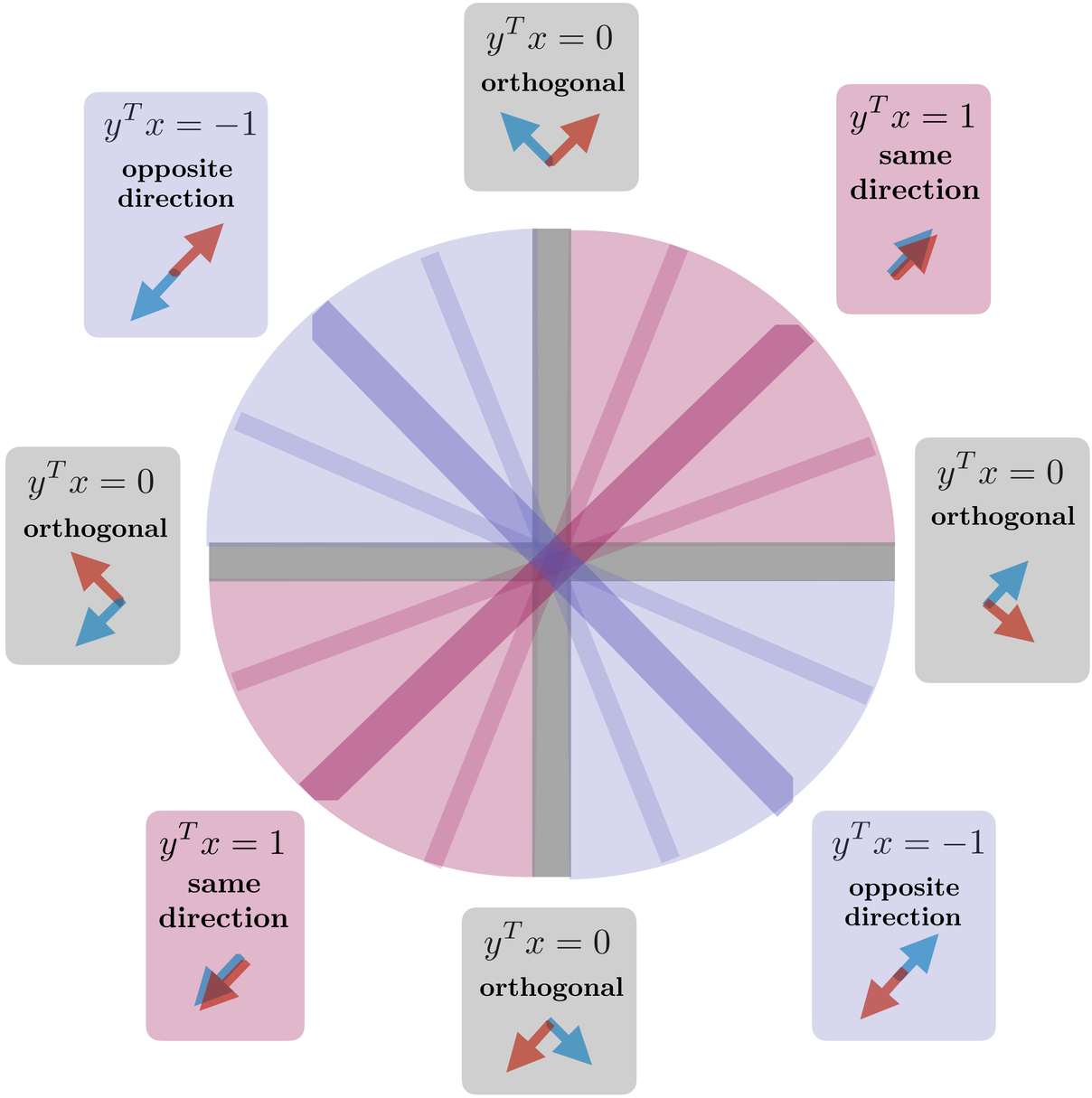


$$y^T x = 0$$



Inner products (in high dimensions)

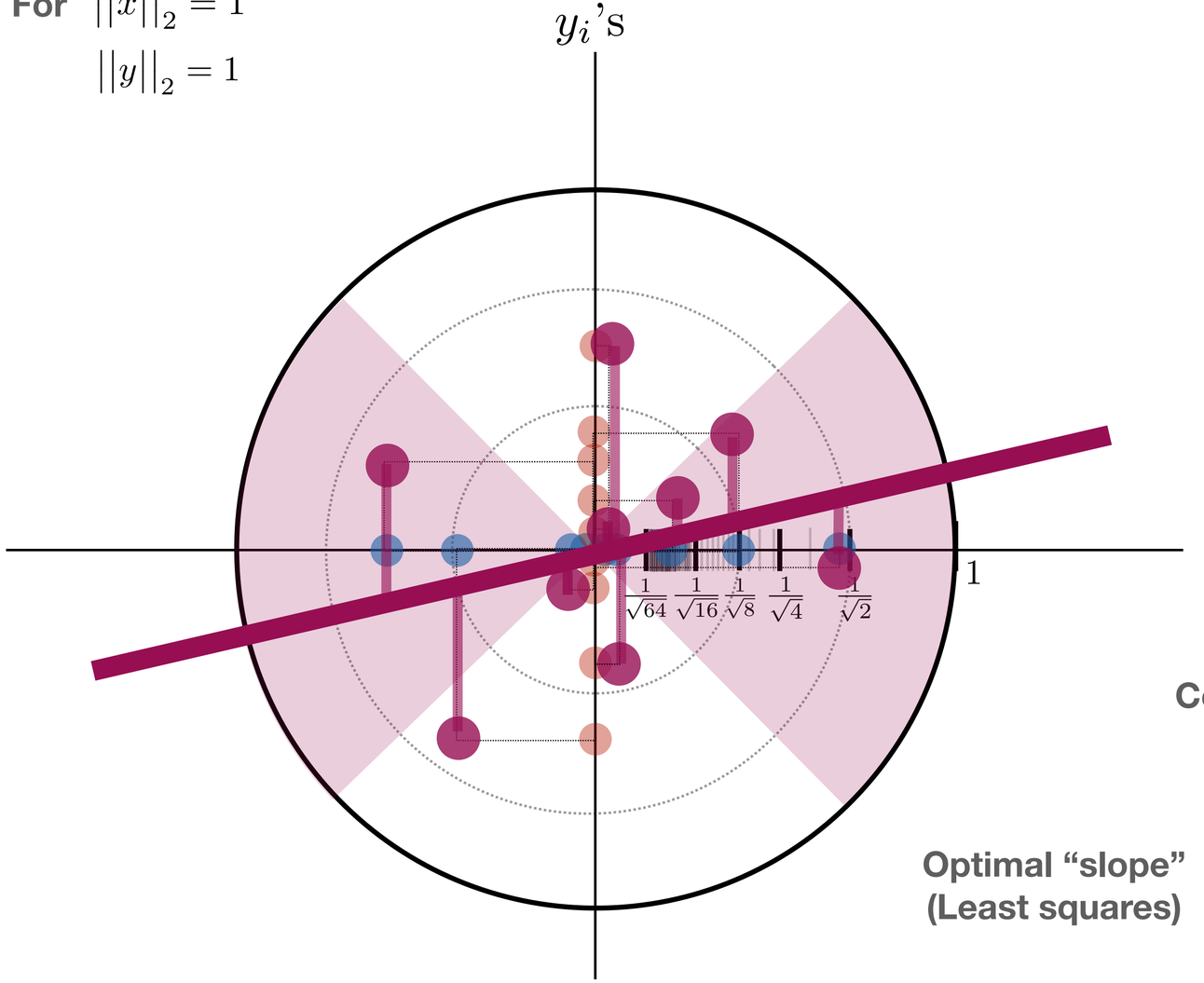


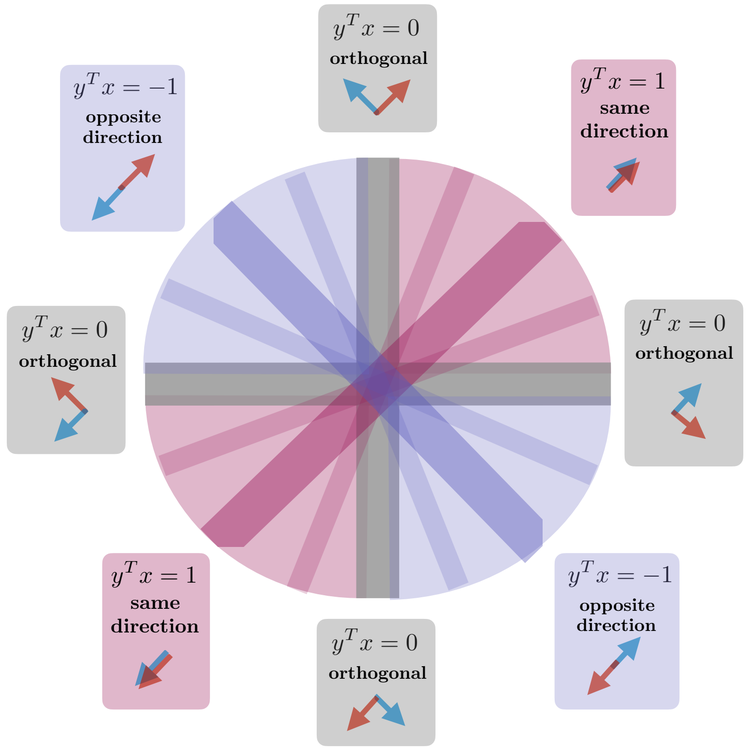


Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





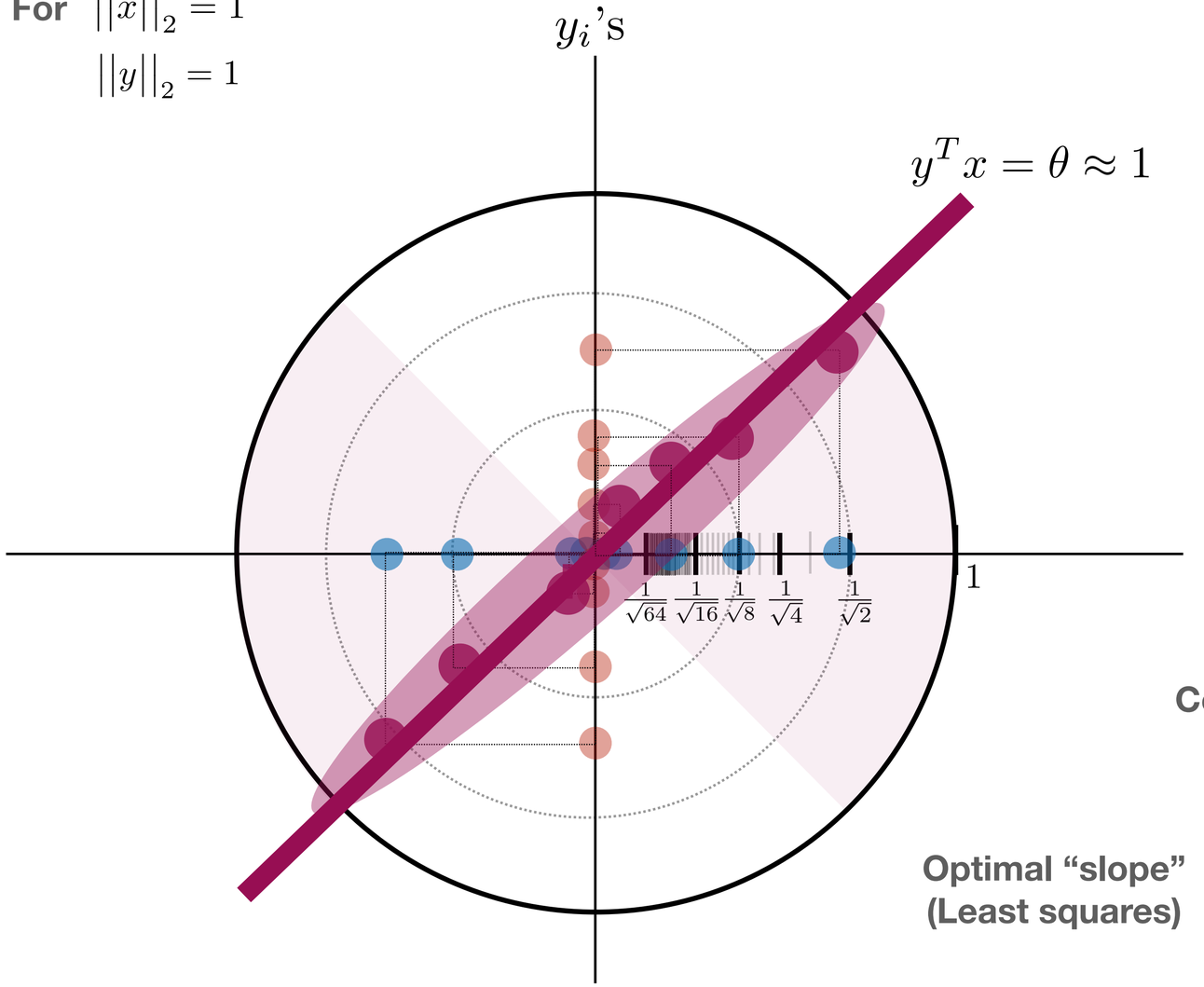
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

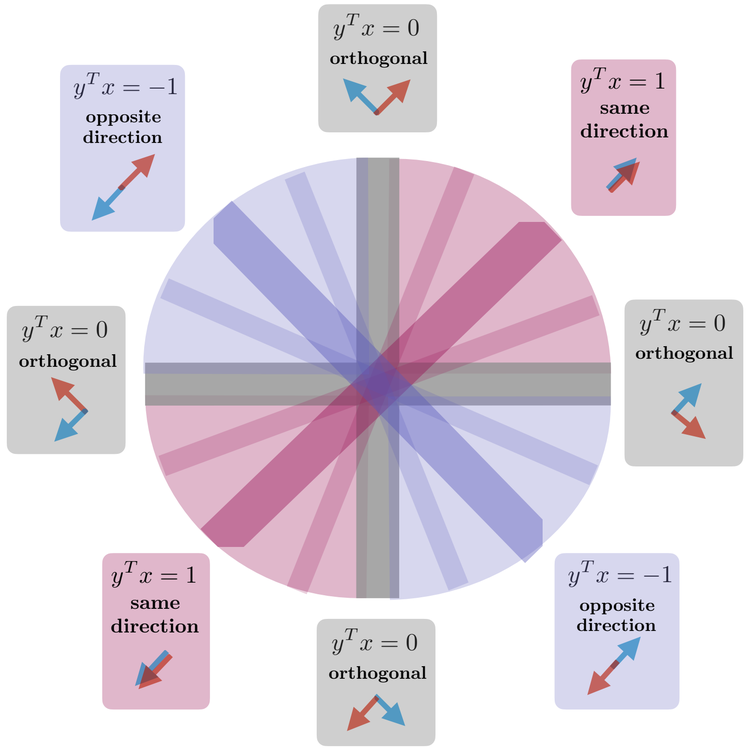
Optimal "slope" (Least squares) $\min_\theta ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





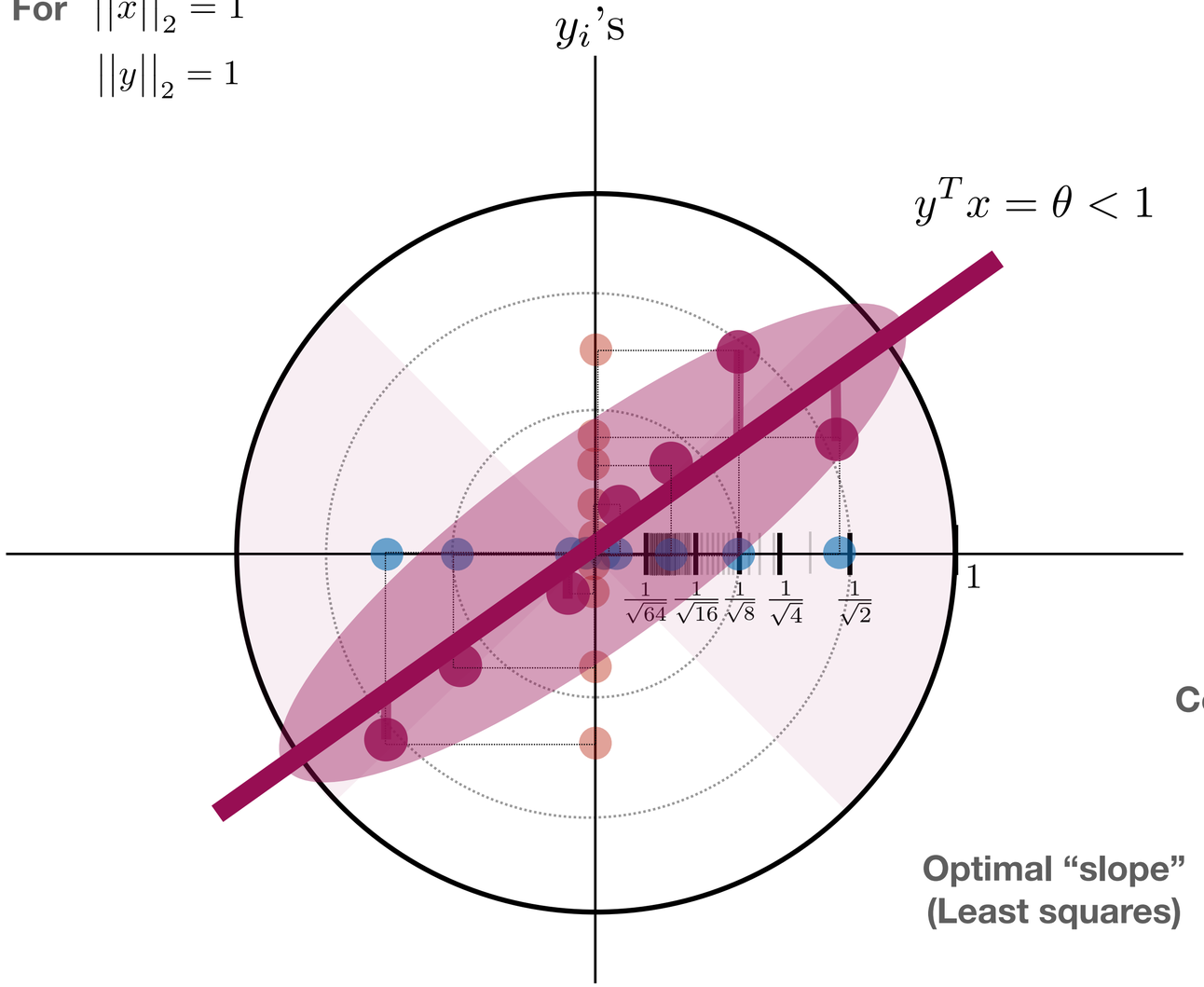
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

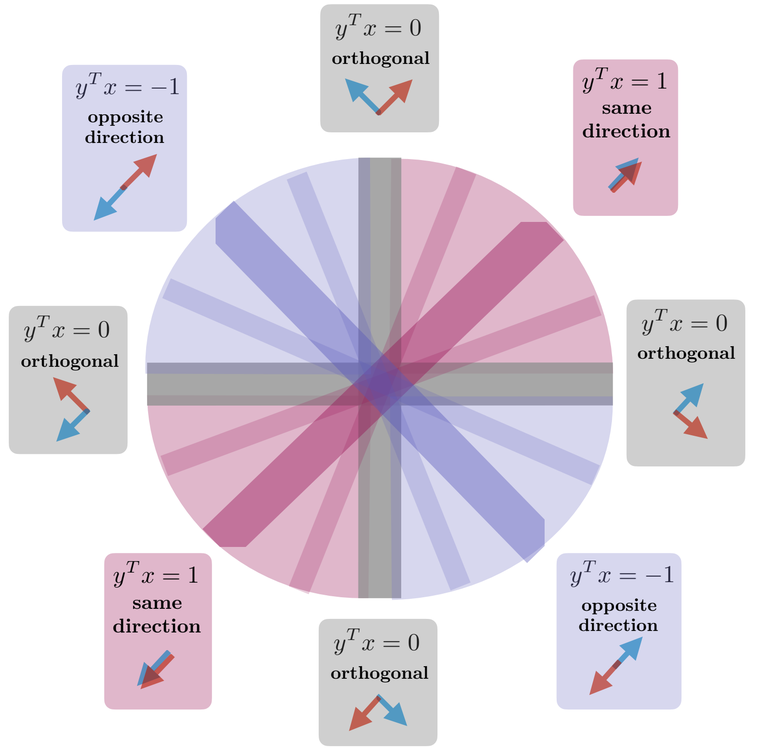
Optimal "slope" (Least squares) $\min_\theta \; ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

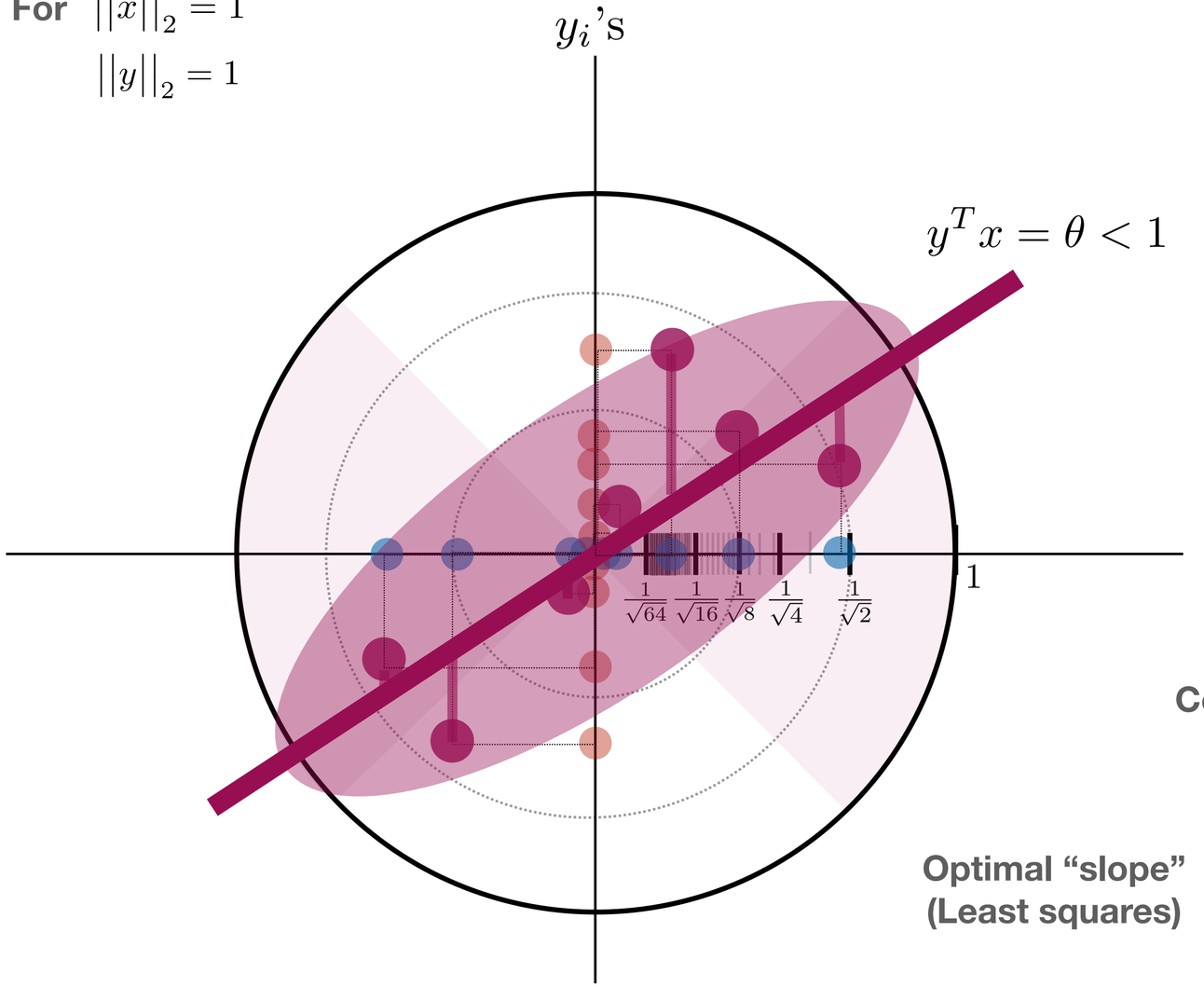
Optimal "slope" (Least squares)

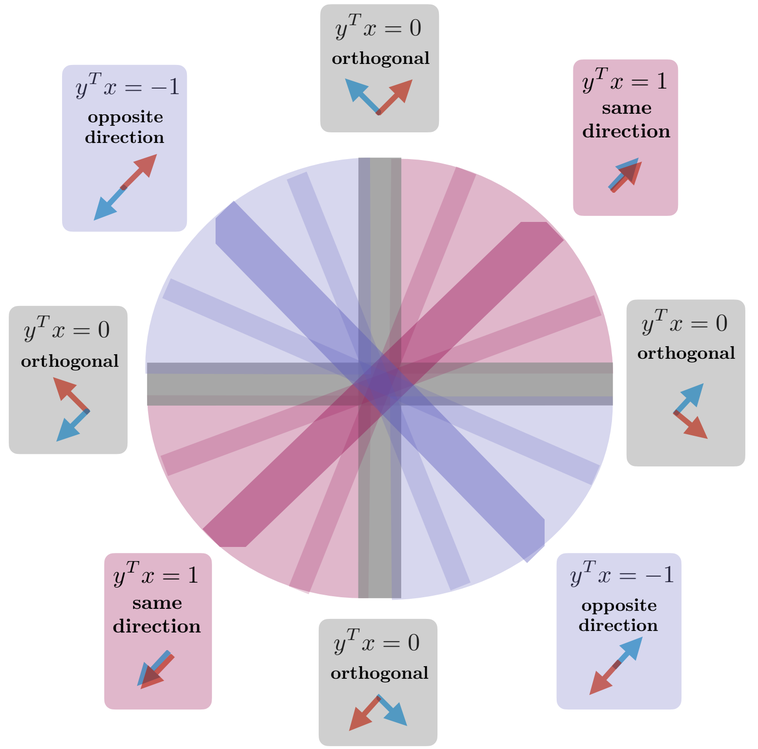
$$\min_\theta \; ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





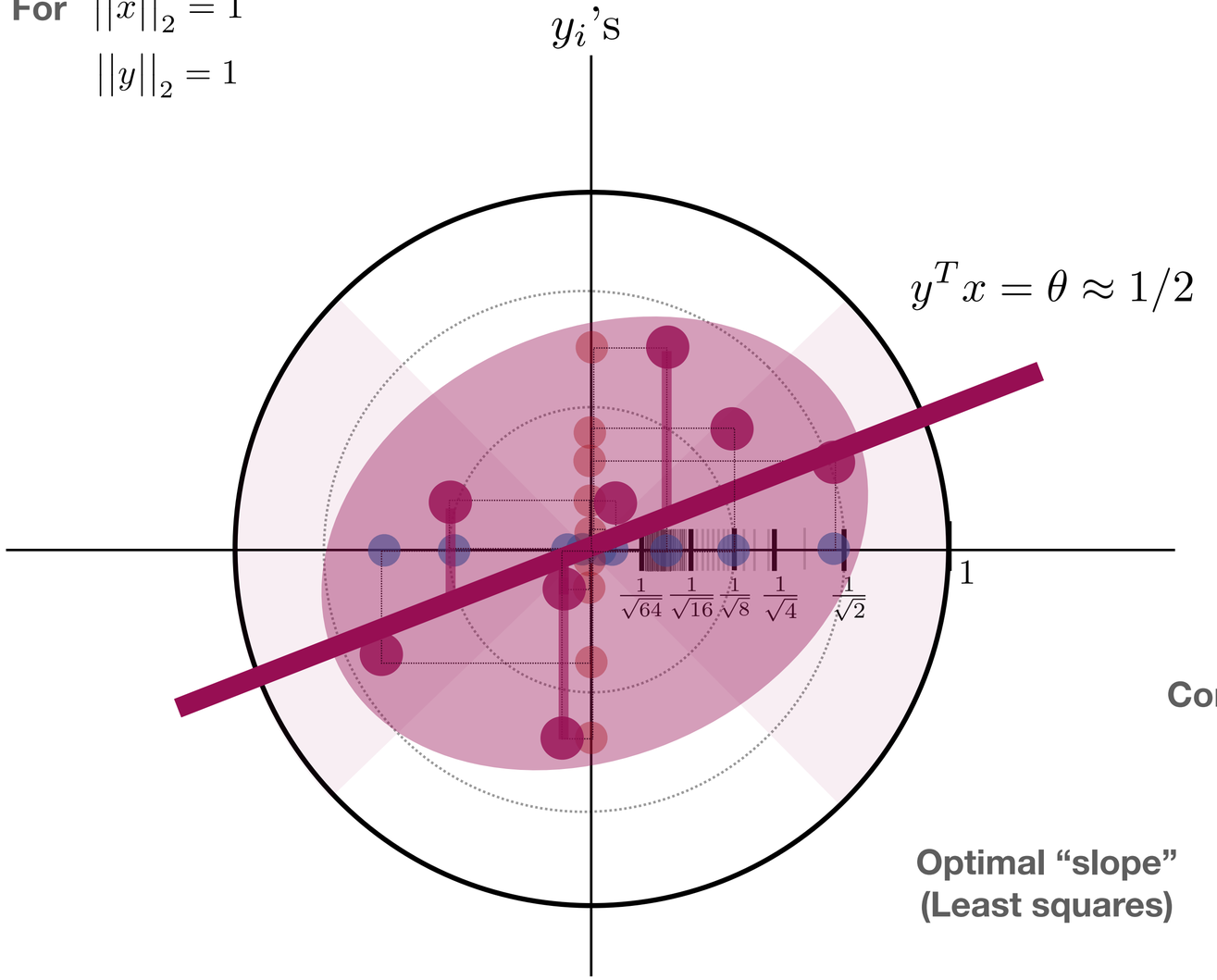
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

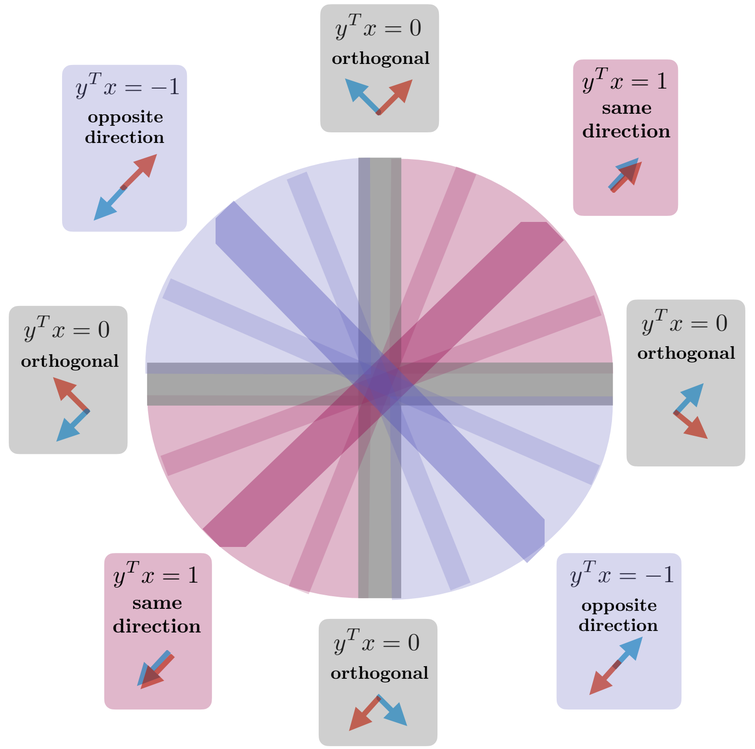
Optimal "slope" (Least squares) $\min_\theta \; ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





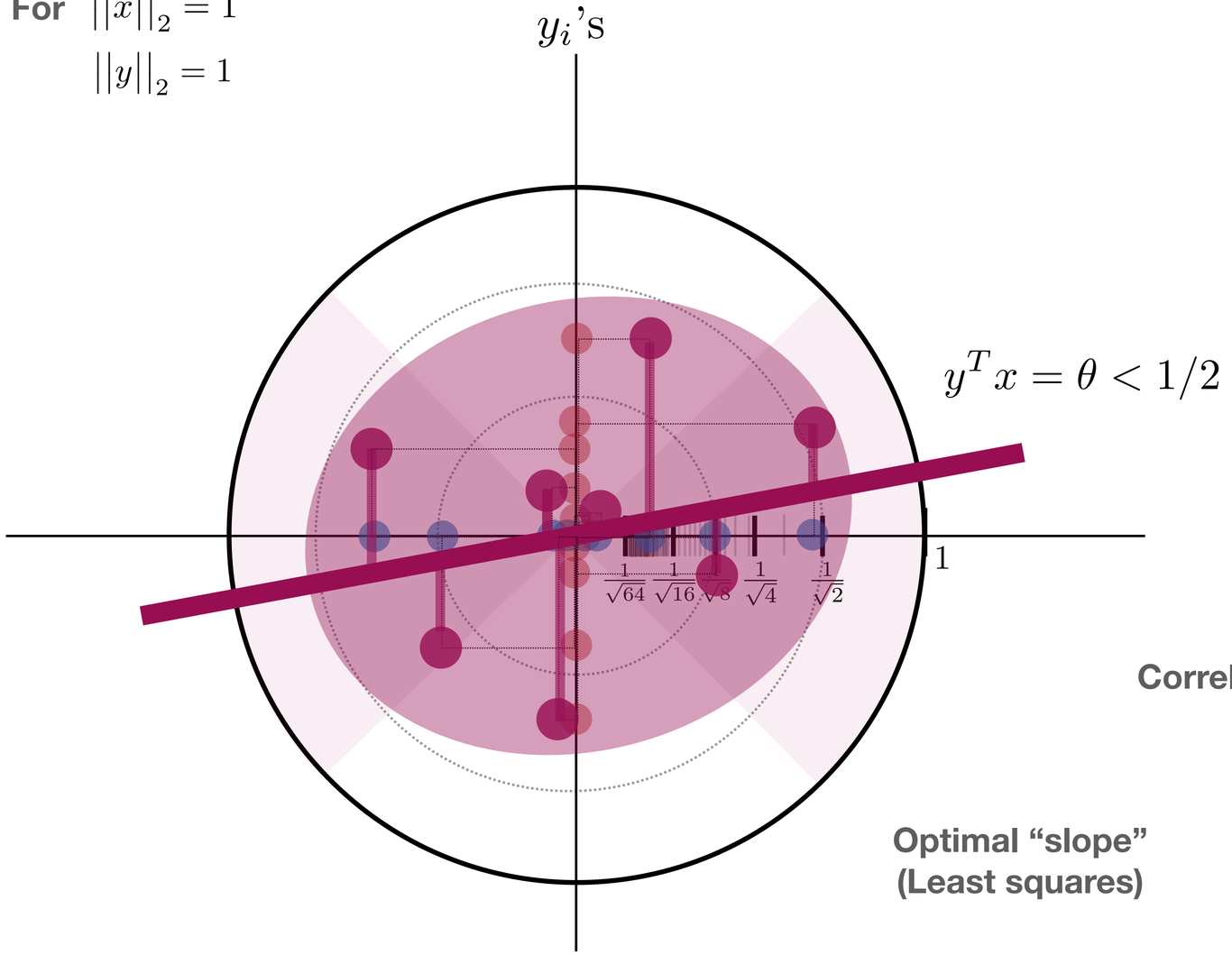
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

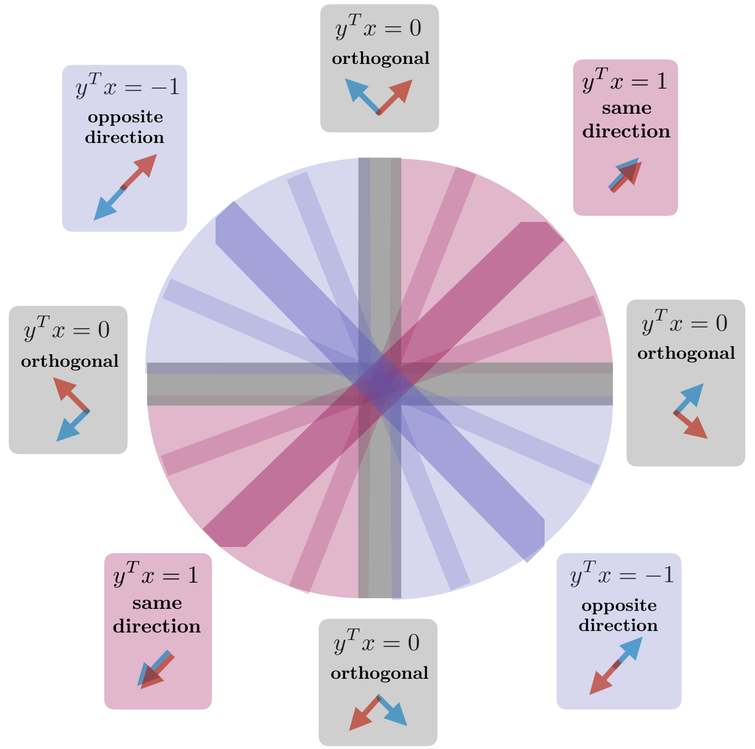
Optimal "slope" (Least squares) $\min_\theta ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





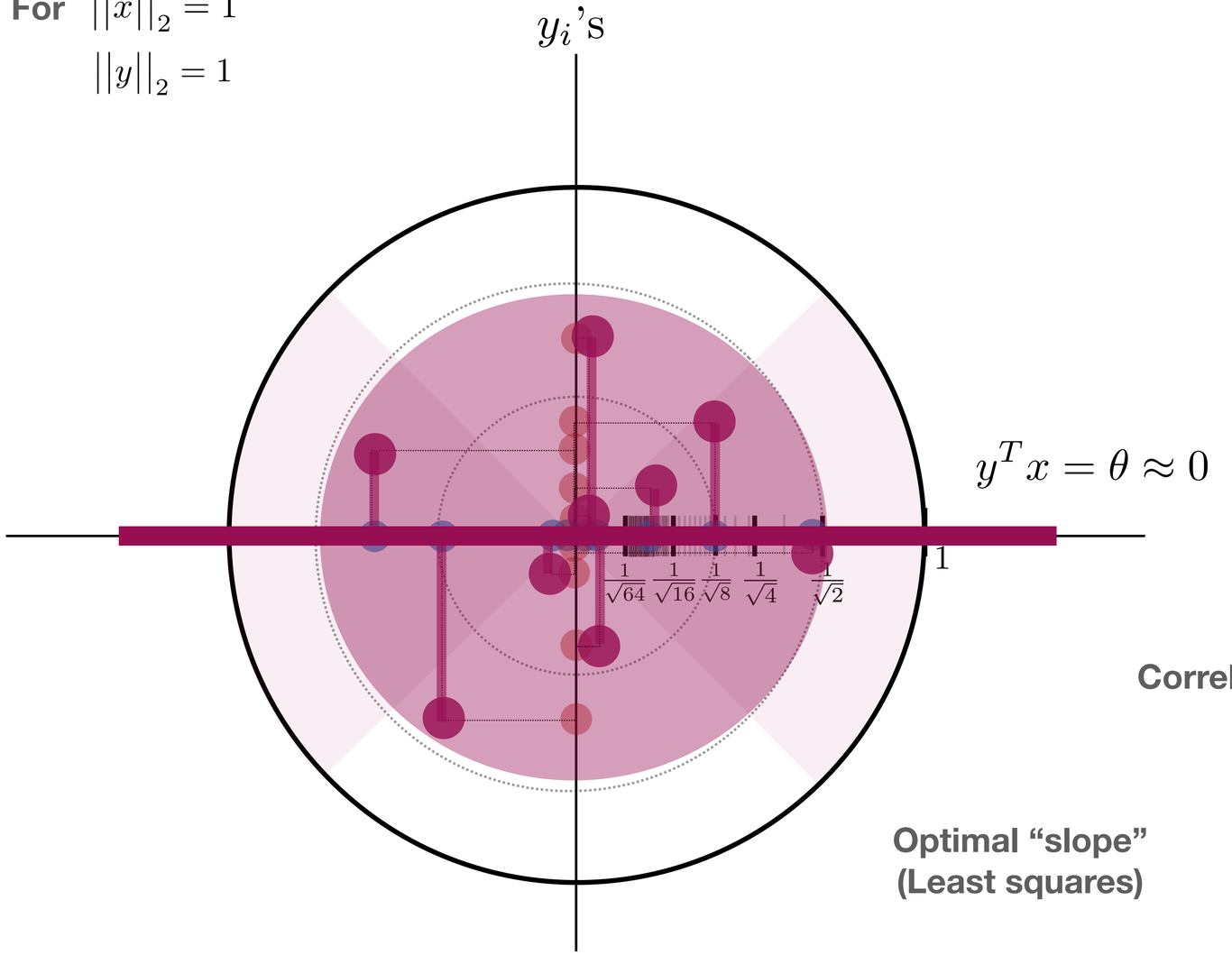
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

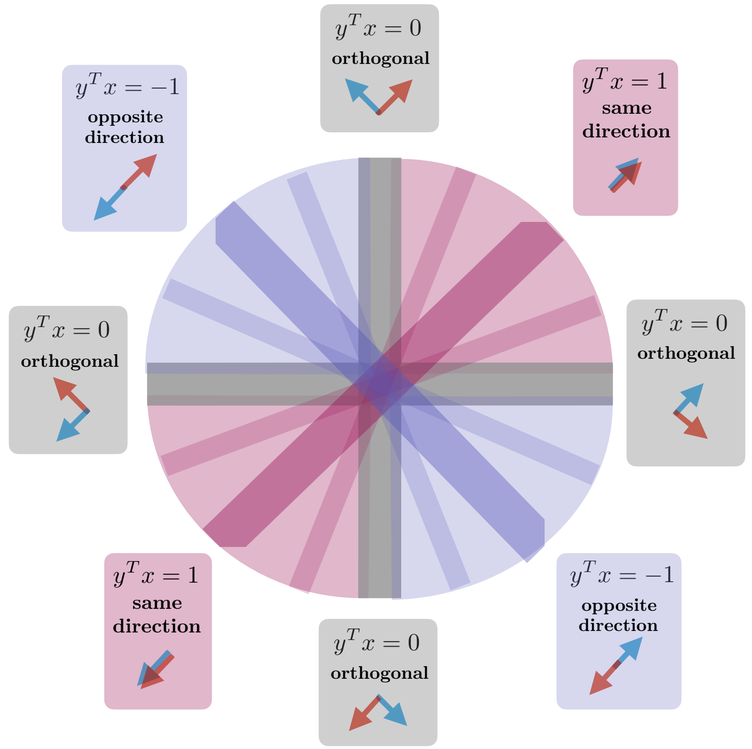
Optimal "slope" (Least squares) $\min_\theta ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





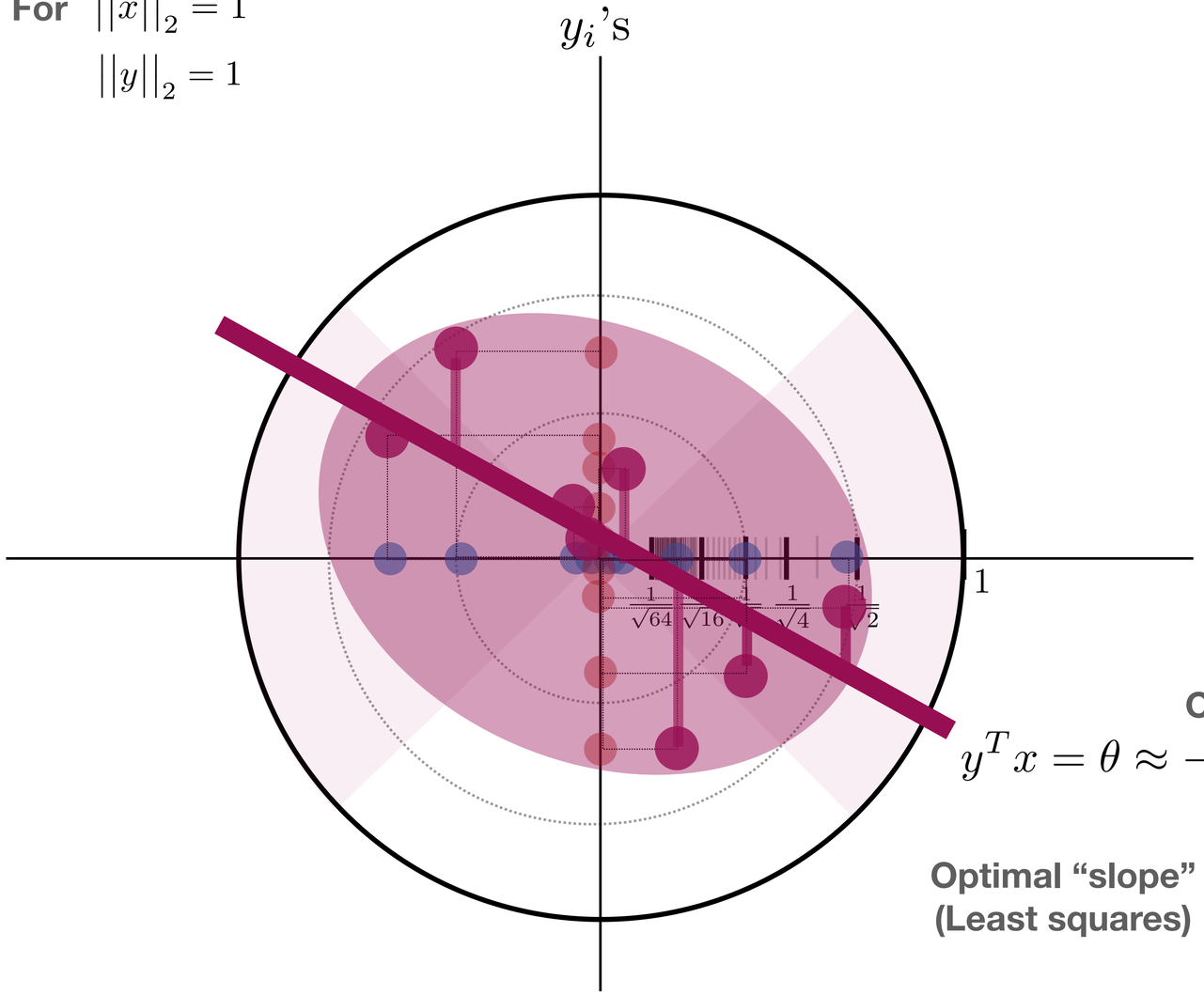
Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

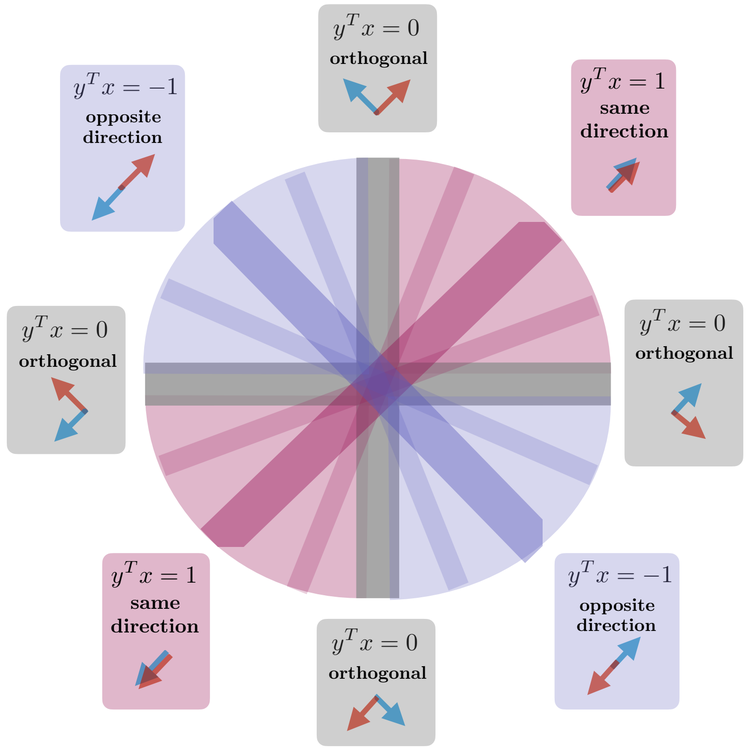
Optimal "slope" (Least squares) $\min_\theta ||y - x\theta||$ $\theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





Correlation

$$y^T x = \theta \approx -1/2$$

$$y^T x = ||y||_2 ||x||_2 \cos\theta$$

Optimal "slope" (Least squares)

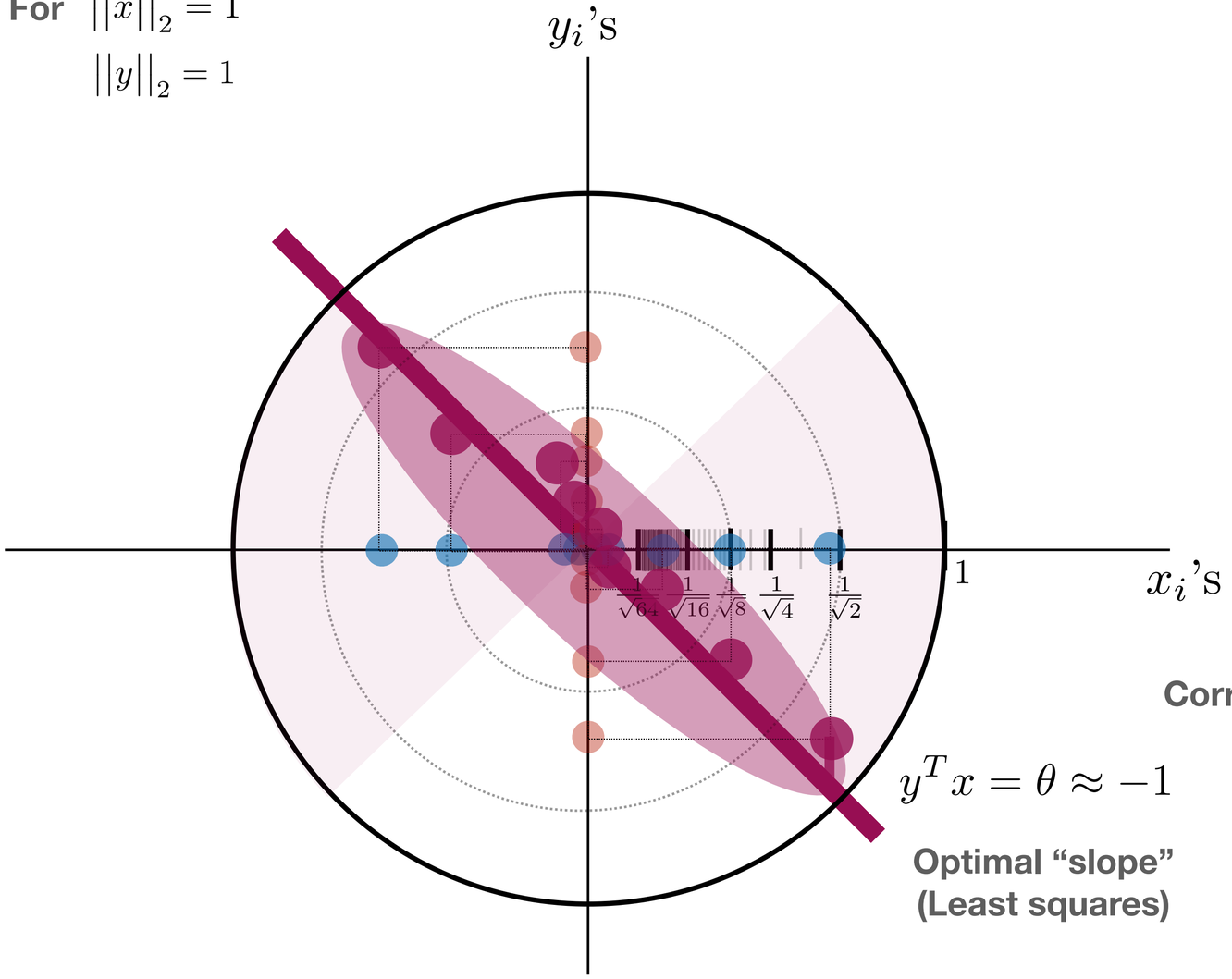
$$\min_\theta \; ||y - x\theta||$$

$$\theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

Relationship to Regression

For $||x||_2 = 1$

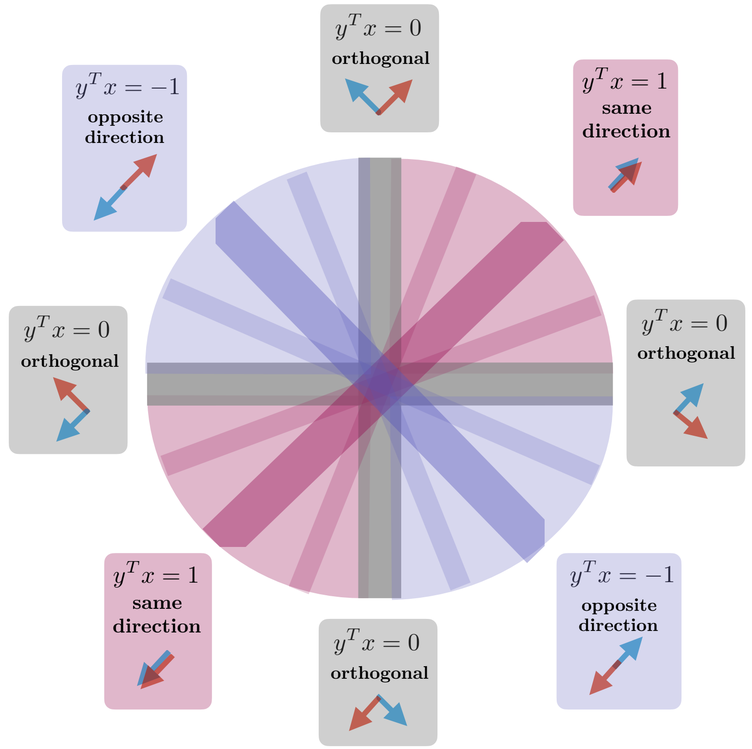
$||y||_2 = 1$



$y^T x = \theta \approx -1$

Optimal "slope" (Least squares)

$$\min_\theta \ ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \ = \frac{||y||_2}{||x||_2} \cos\theta$$

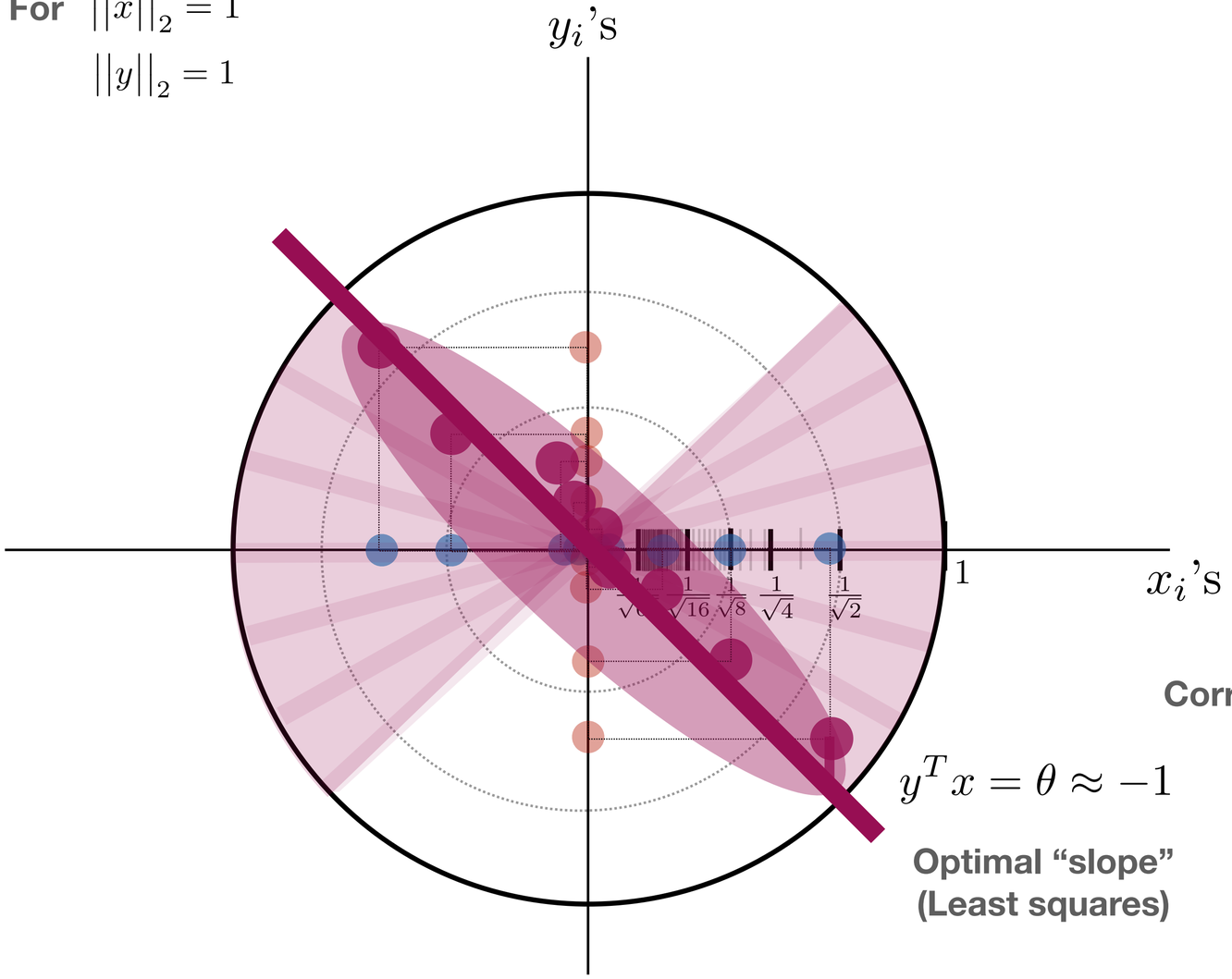


Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

Relationship to Regression

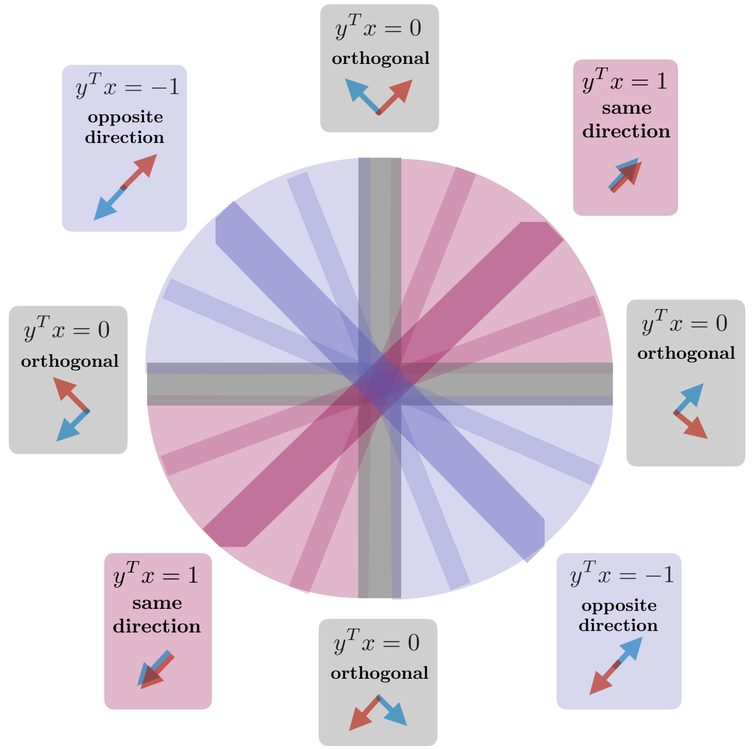
For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope" (Least squares)

$$\min_\theta \; ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

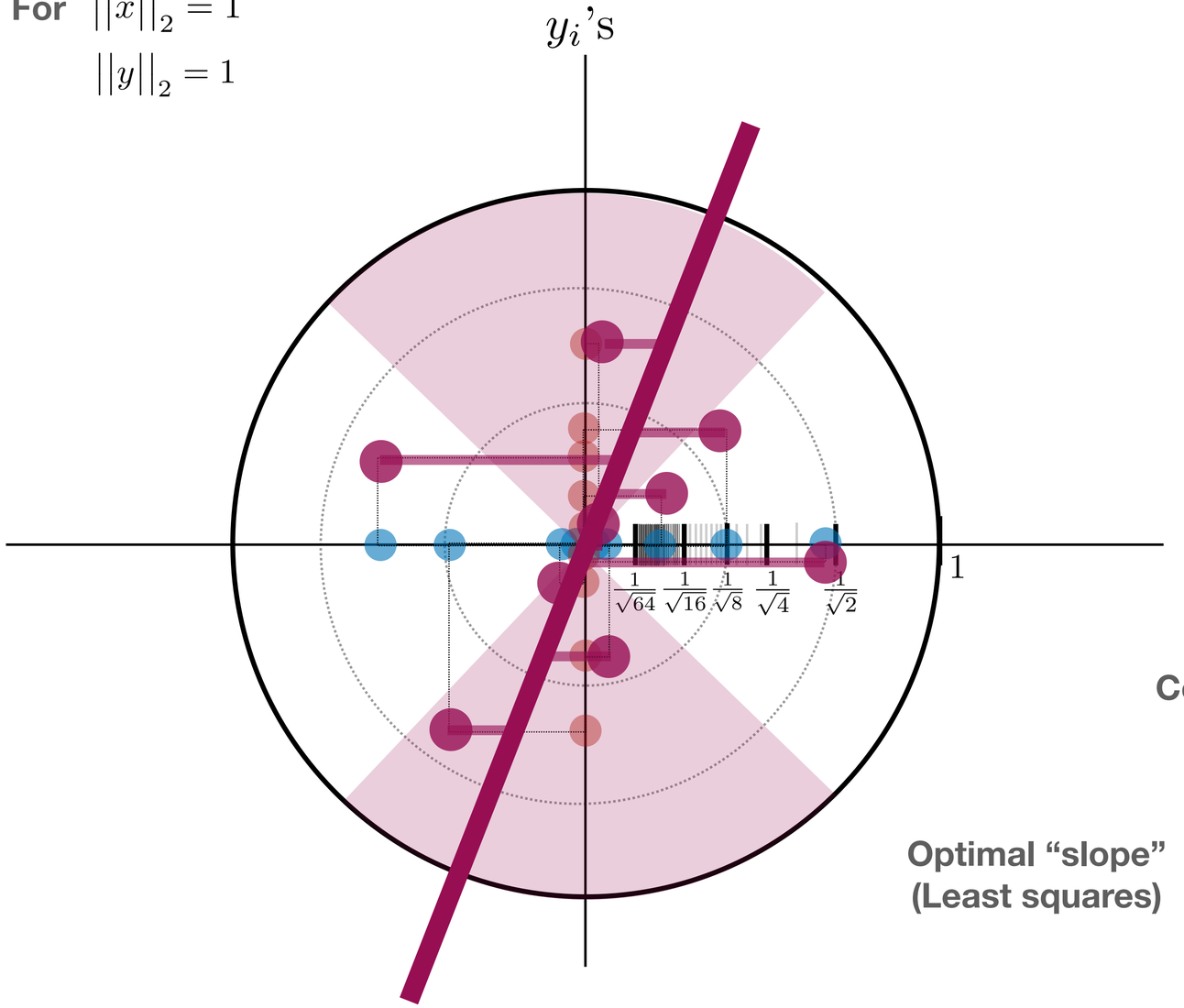


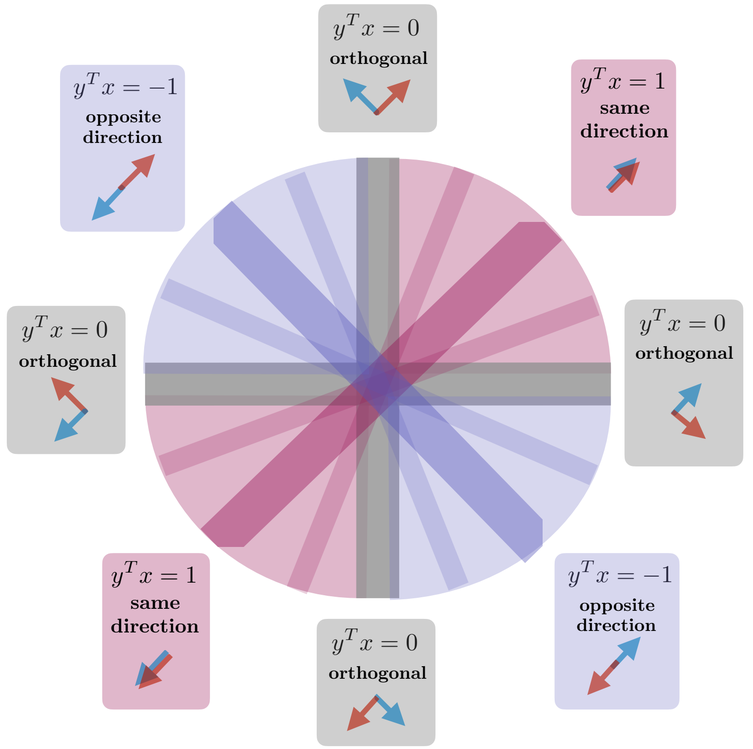
Correlation $$y^T x = ||y||_2 ||x||_2 \cos\theta$$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





Correlation

$$y^T x = ||y||_2 ||x||_2 \cos\theta$$

Optimal "slope" (Least squares)

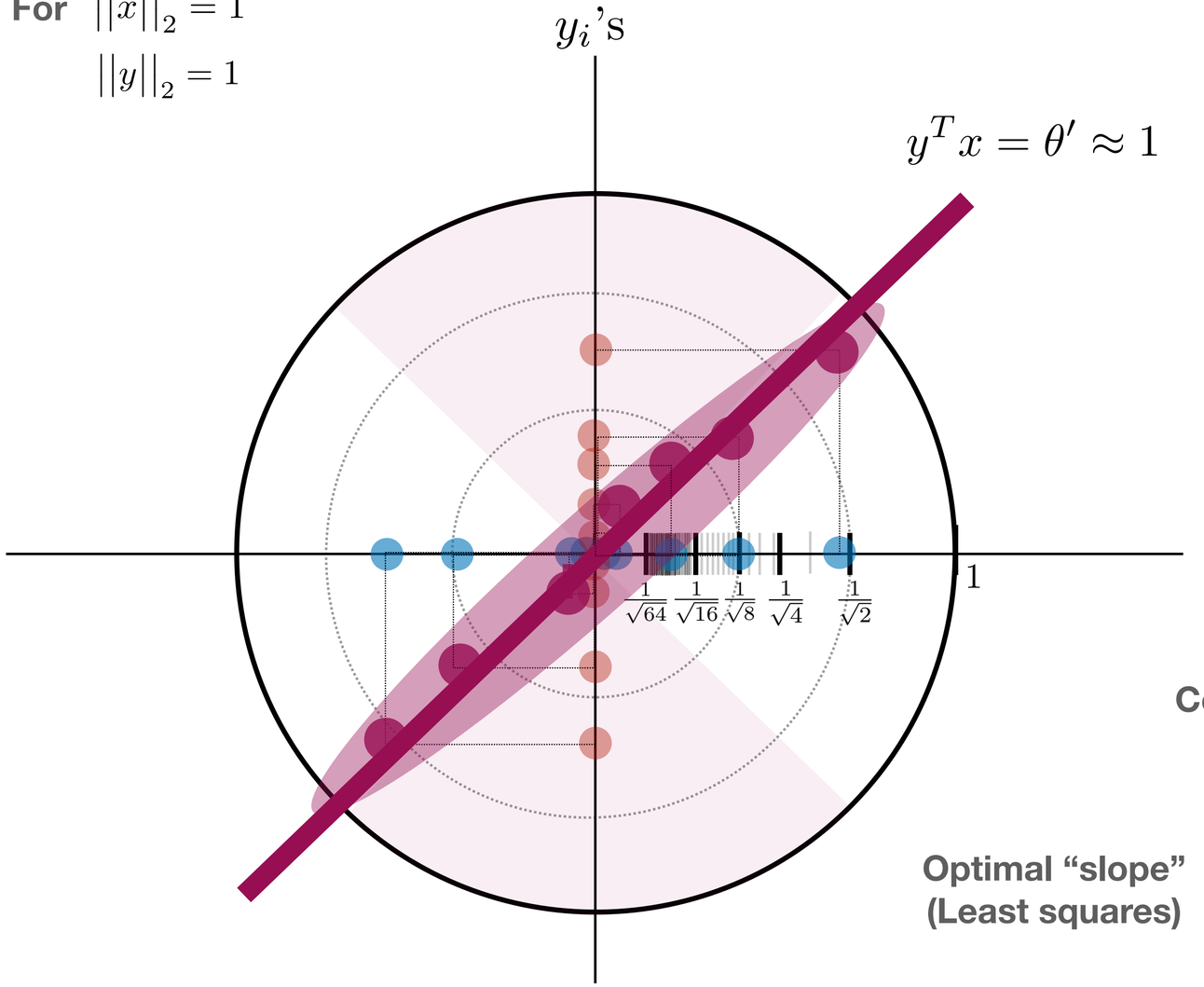
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

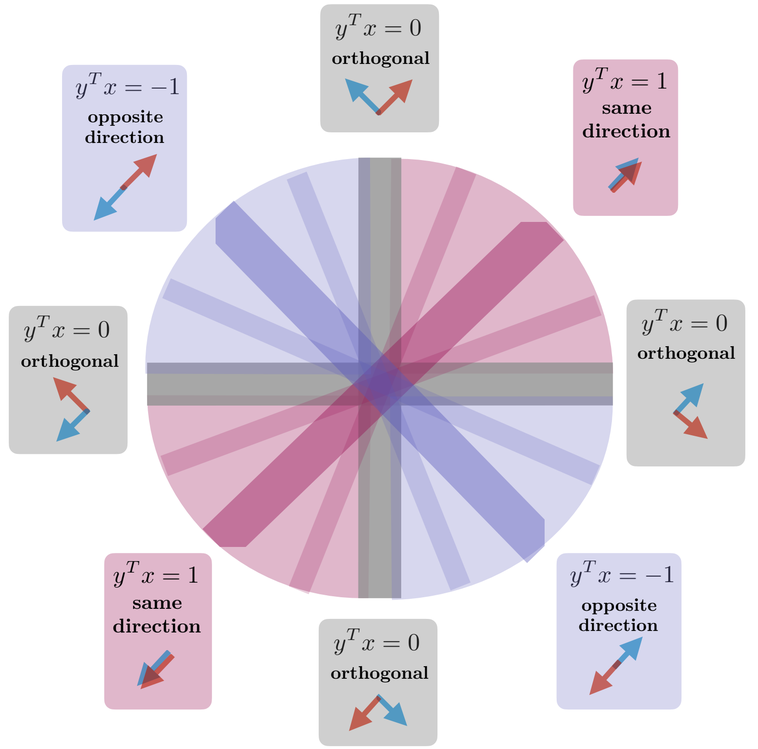
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope"
(Least squares)



Correlation $$y^T x = ||y||_2 ||x||_2 \cos\theta$$

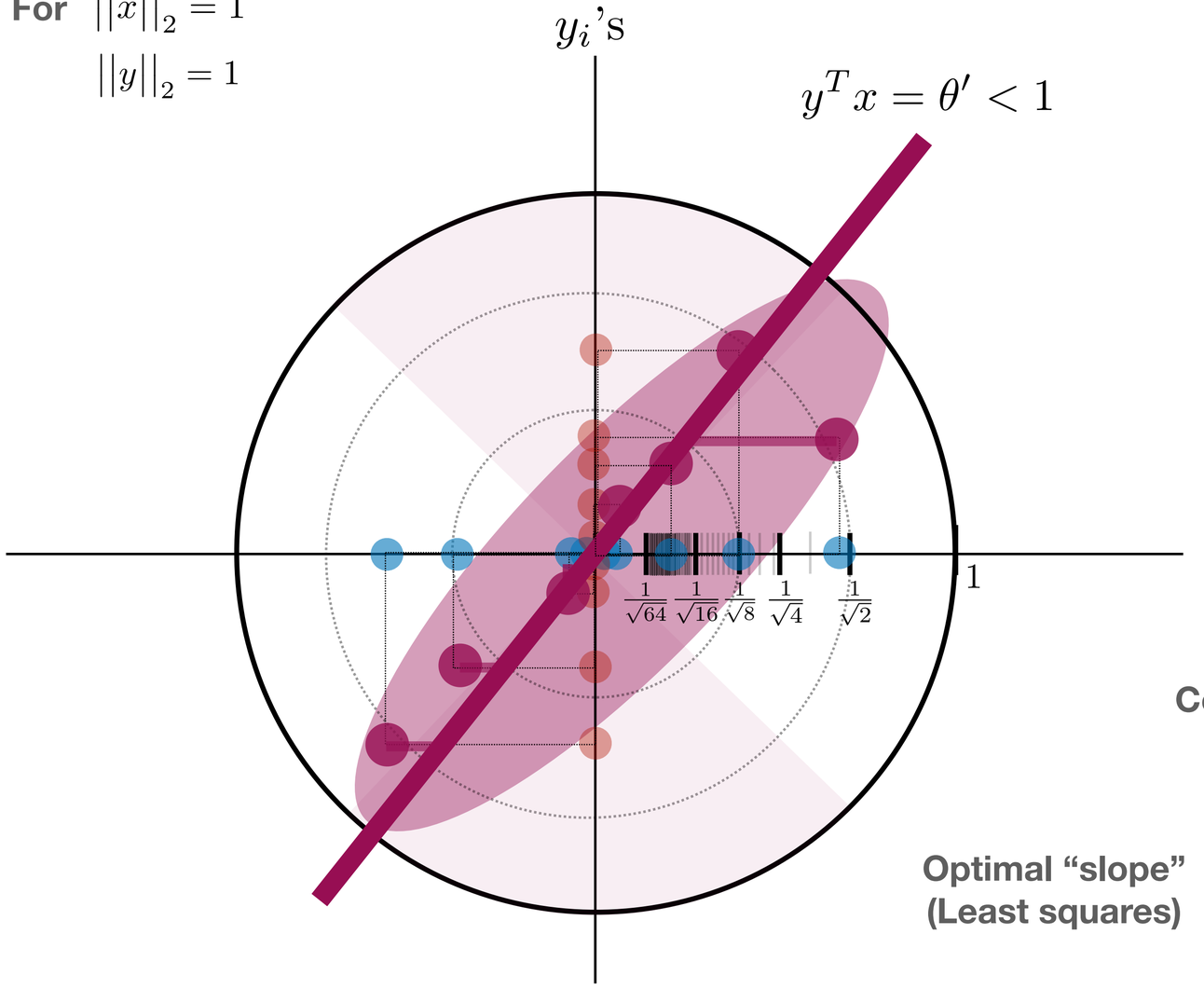
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

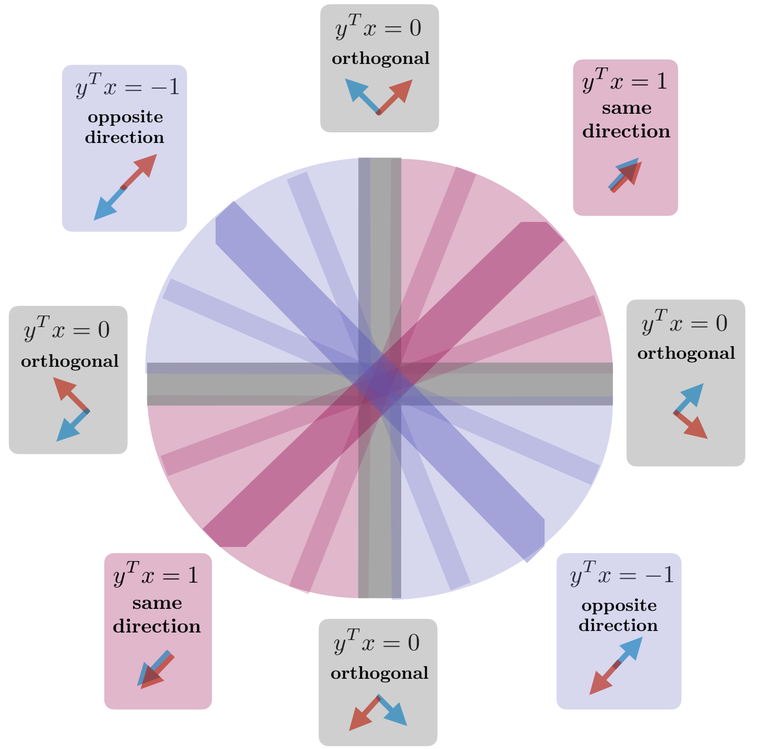
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope"
(Least squares)



Correlation $$y^T x = ||y||_2 ||x||_2 \cos\theta$$

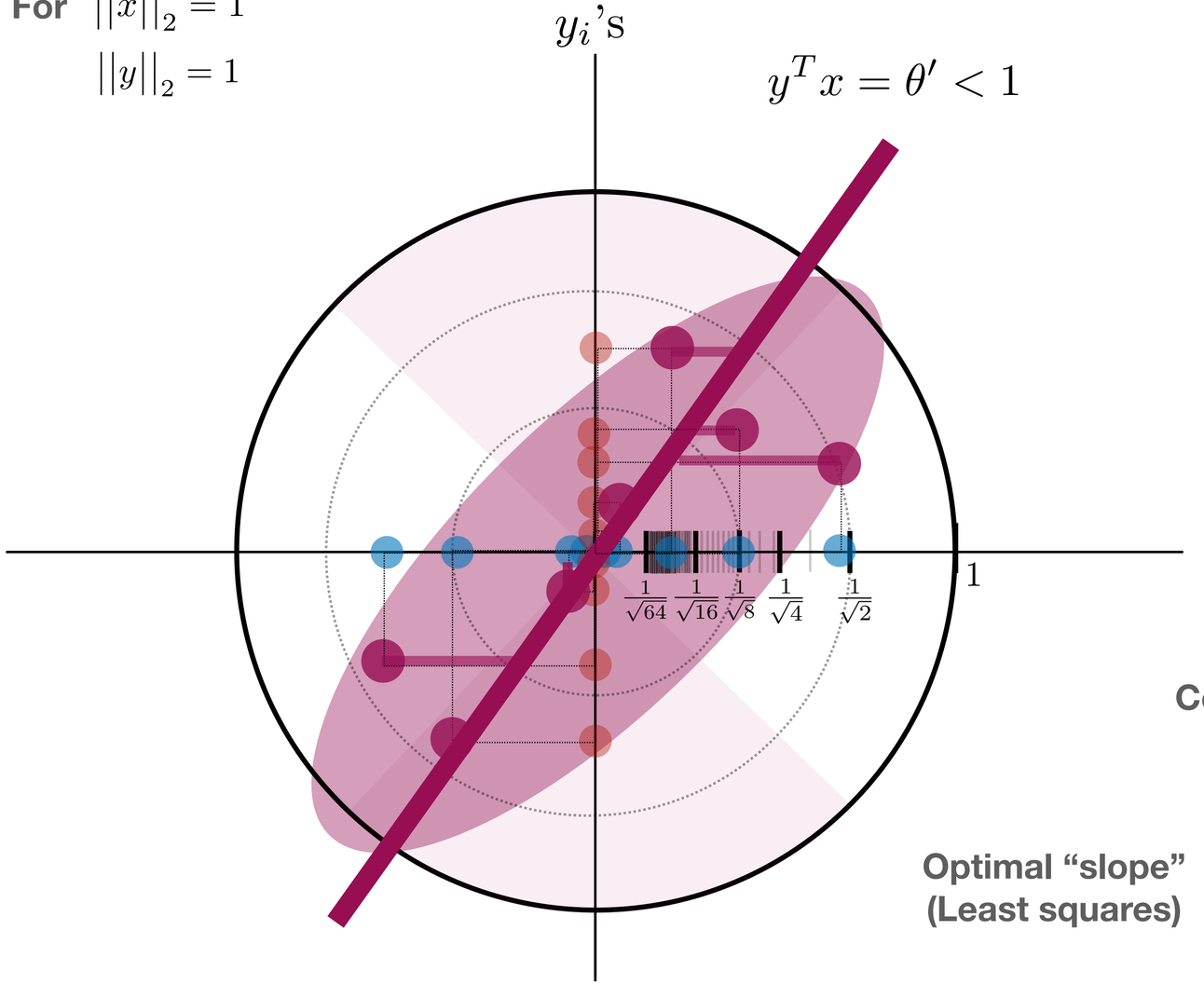
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

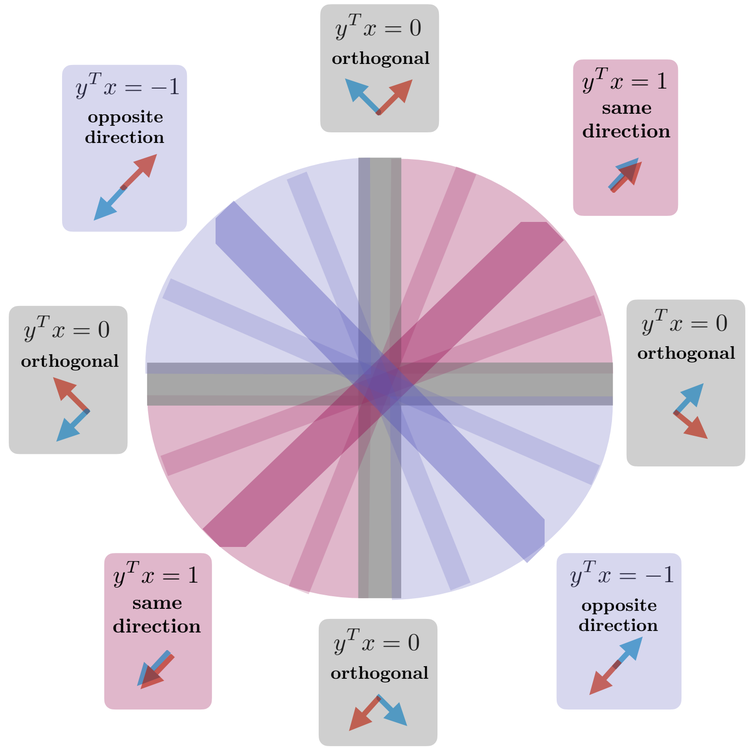
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope" (Least squares)



Correlation $$y^T x = ||y||_2 ||x||_2 \cos\theta$$

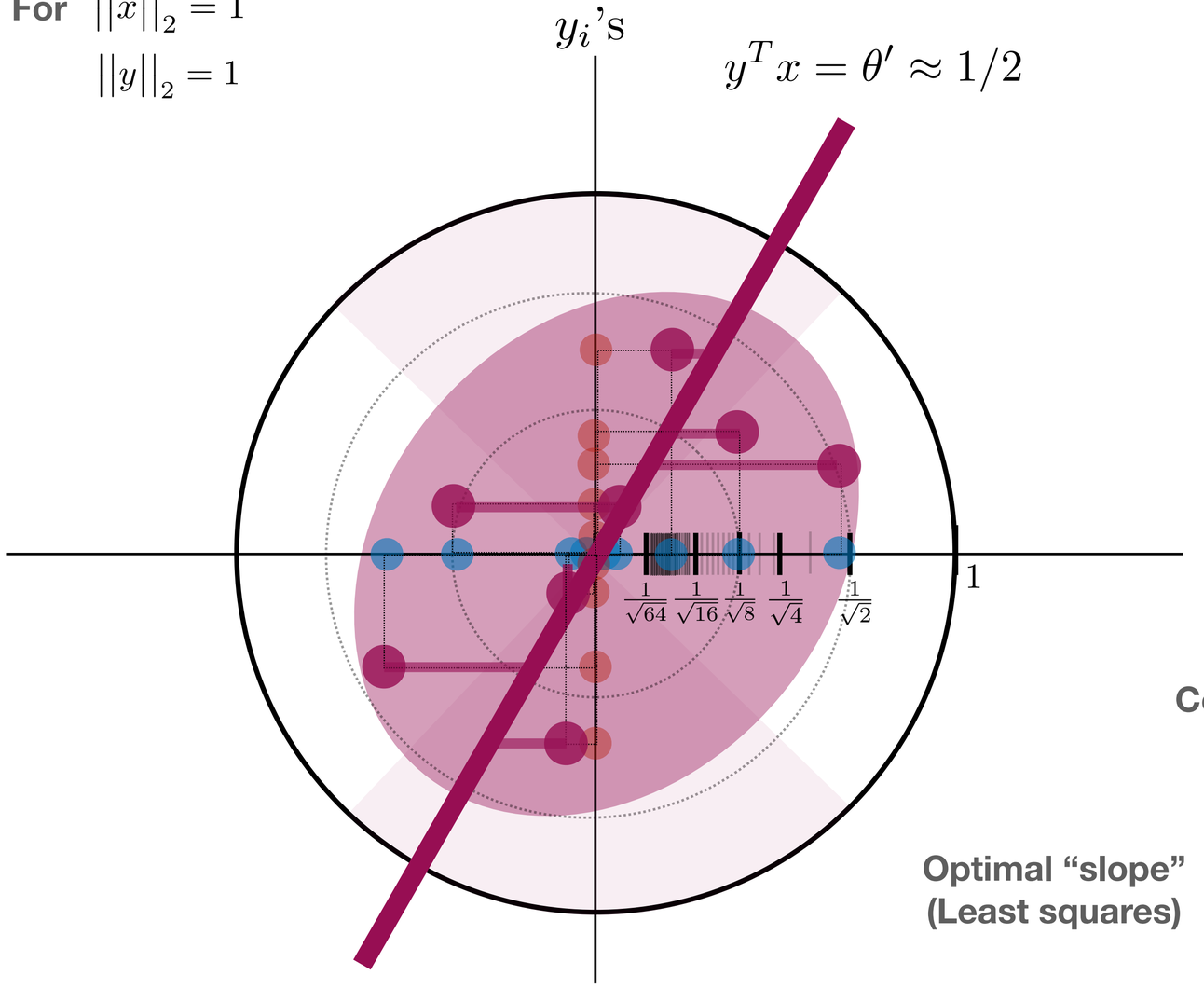
$$\min_\theta \; ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

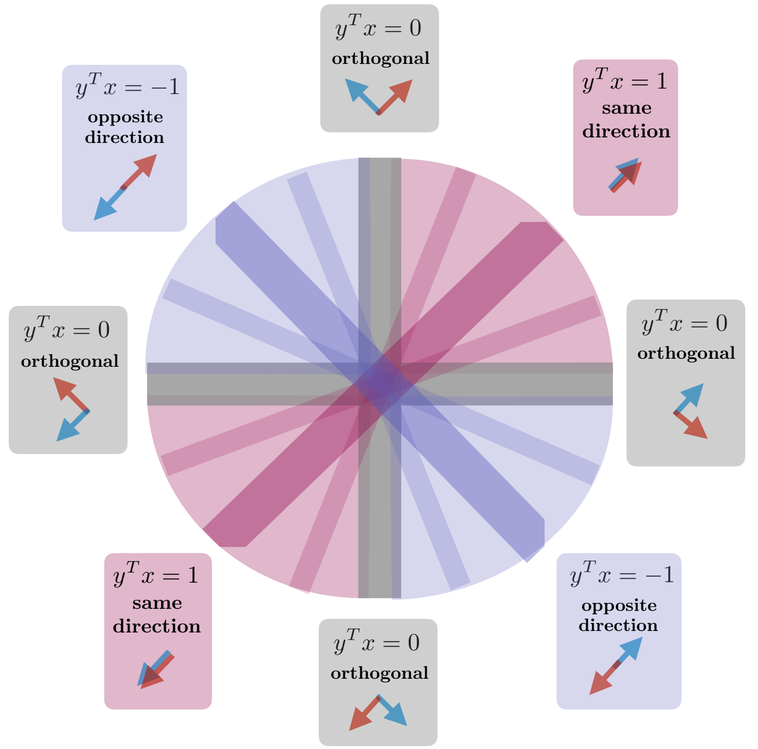
$$\min_{\theta'} \; ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x \; = \frac{||x||_2}{||y||_2} \cos\theta$$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

Optimal "slope" (Least squares)

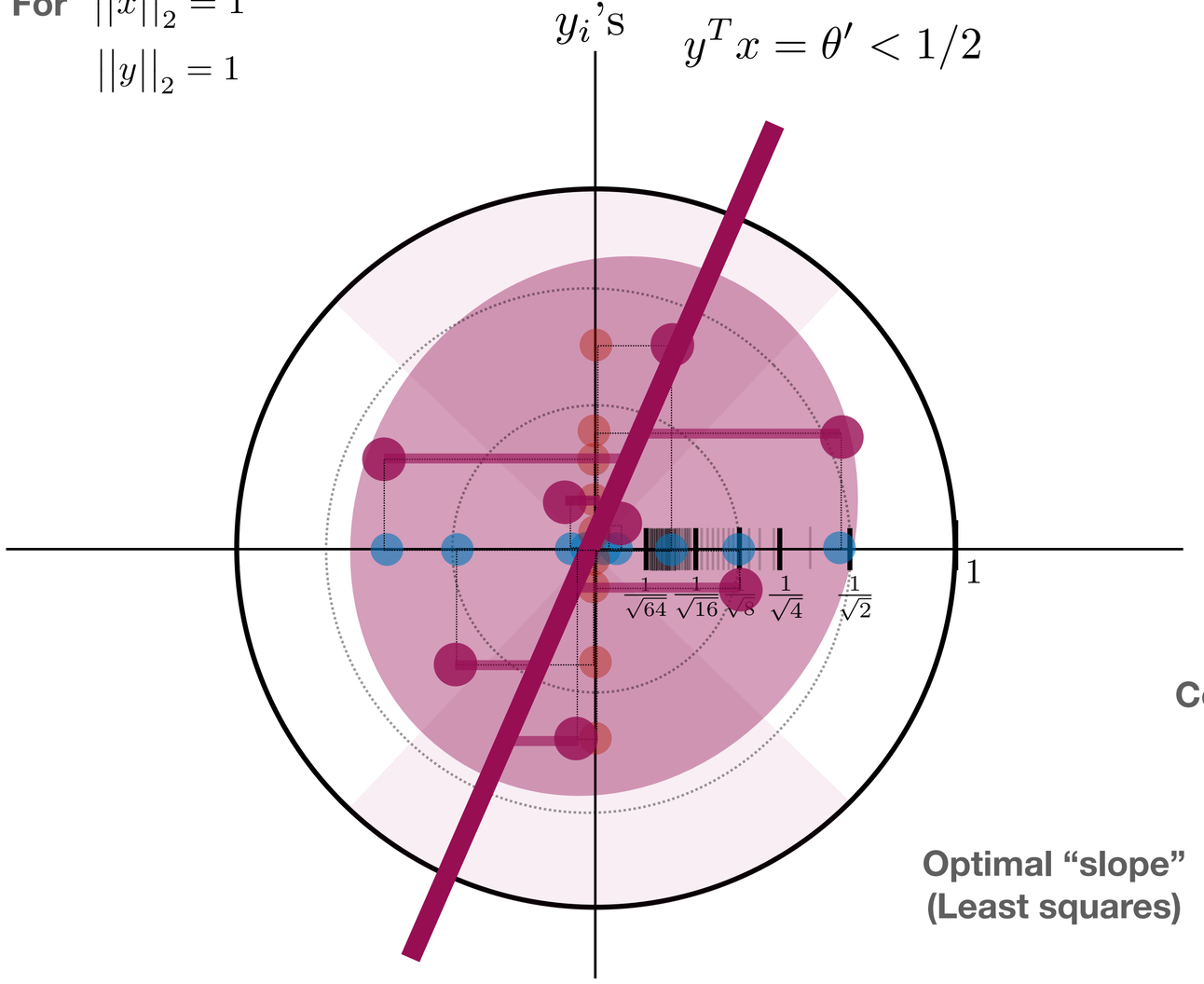
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

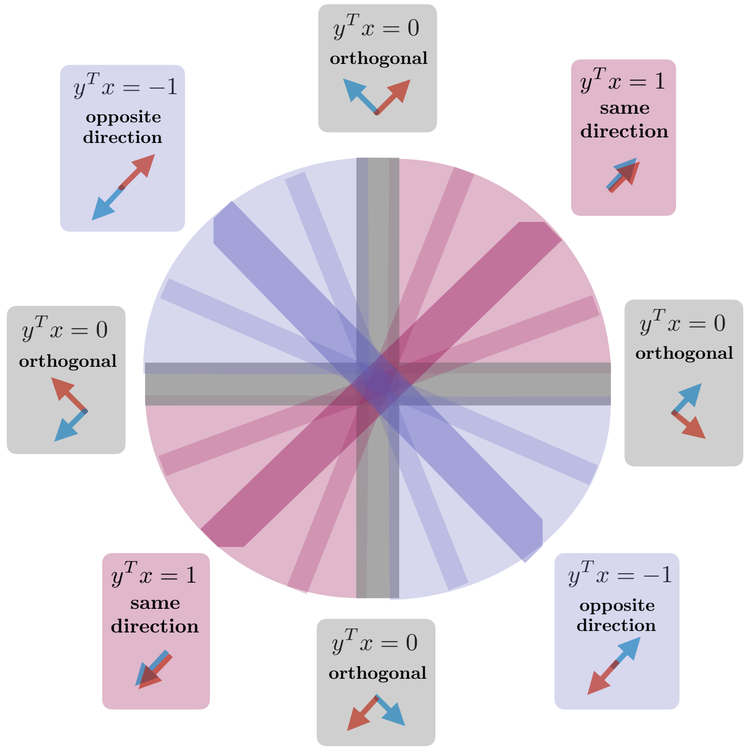
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope"
(Least squares)



Correlation $$y^T x = ||y||_2 ||x||_2 \cos\theta$$

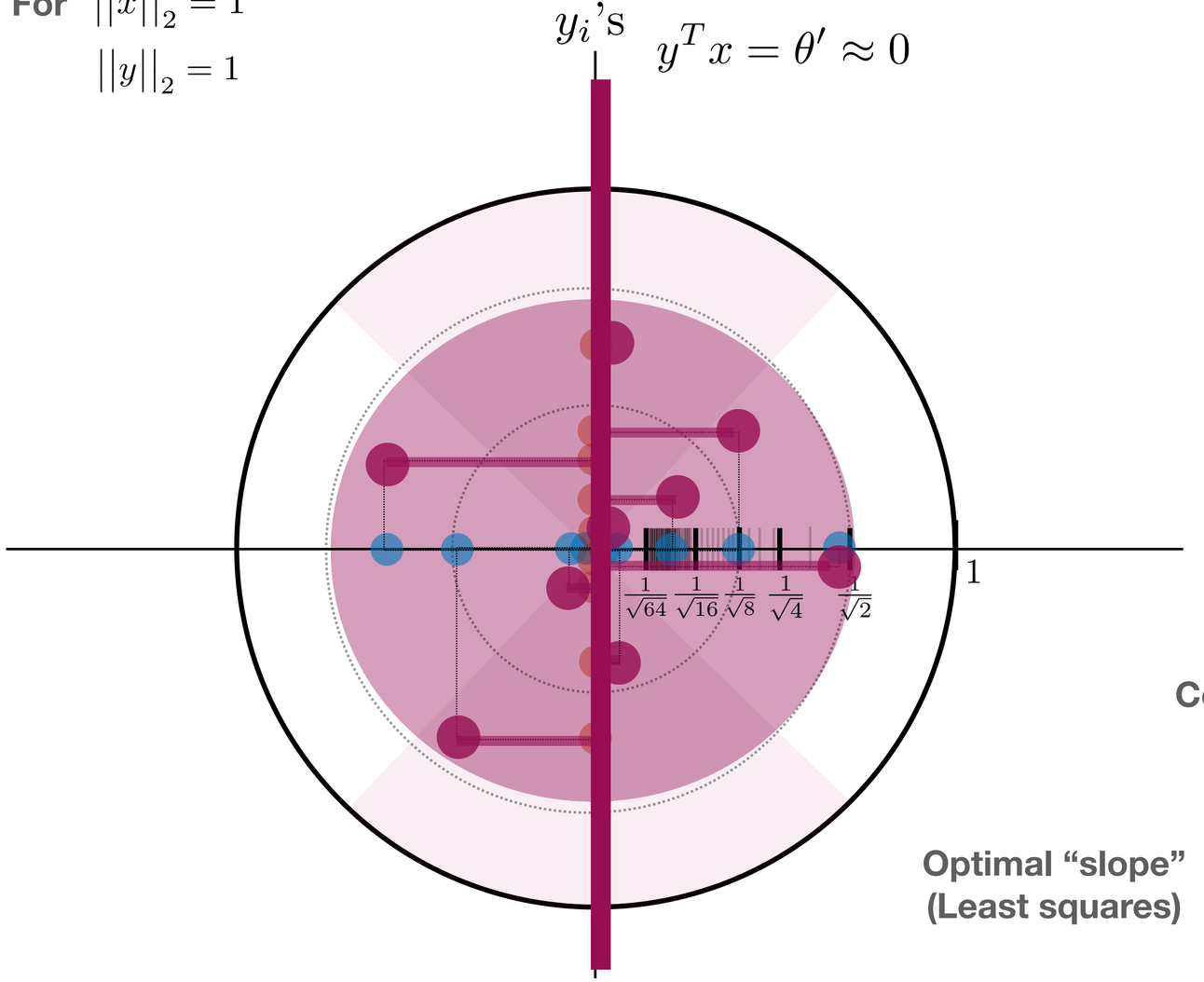
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

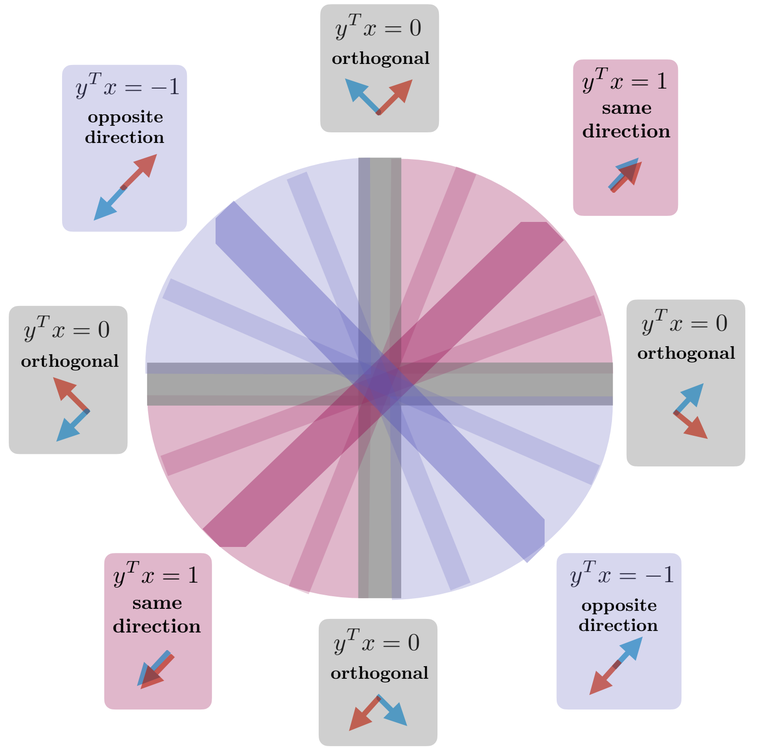
$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$





Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

Optimal "slope" (Least squares)

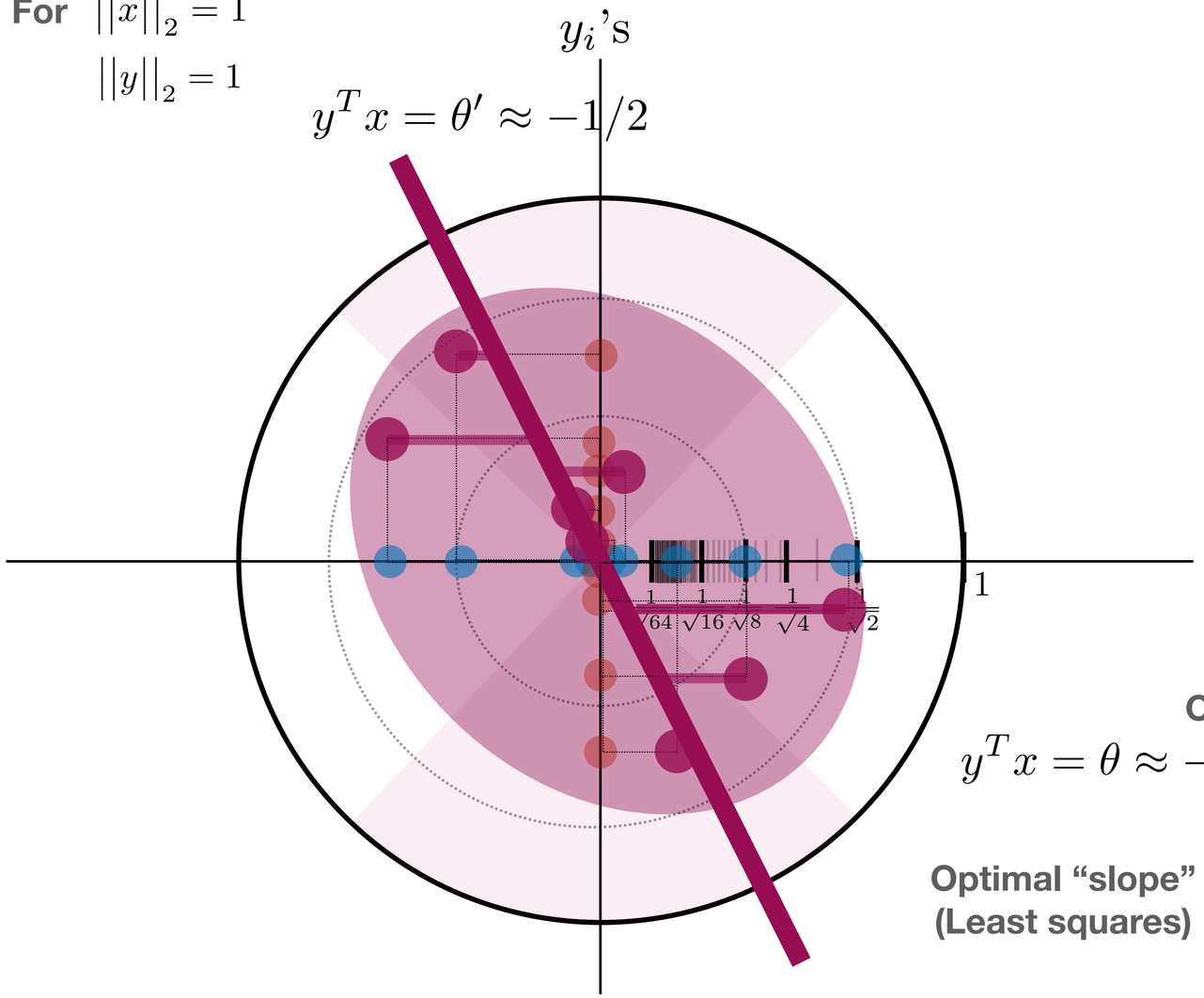
$$\min_\theta \; ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} \; ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x \; = \frac{||x||_2}{||y||_2} \cos\theta$$

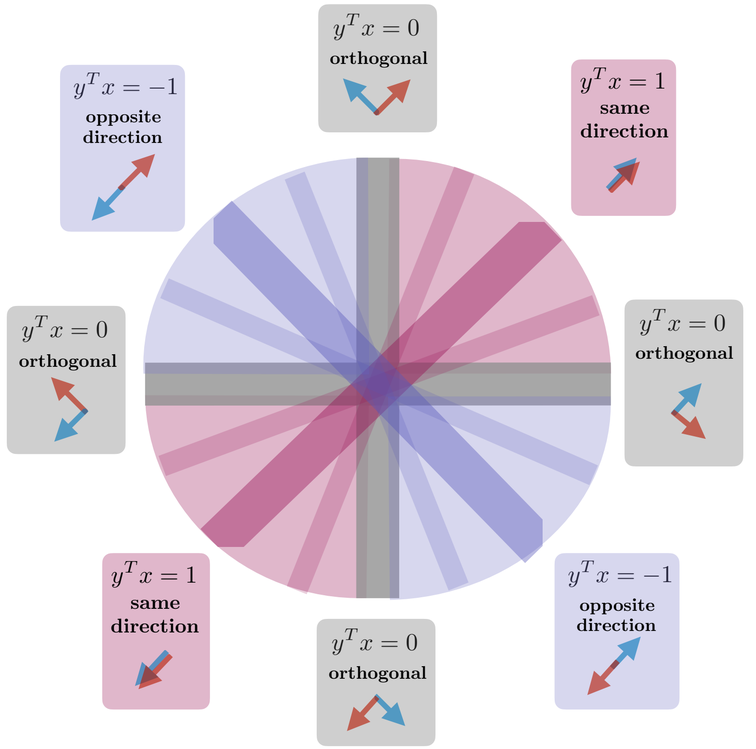
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



$y^T x = \theta \approx -1/2$



Correlation

$$y^T x = ||y||_2 ||x||_2 \cos\theta$$

Optimal "slope" (Least squares)

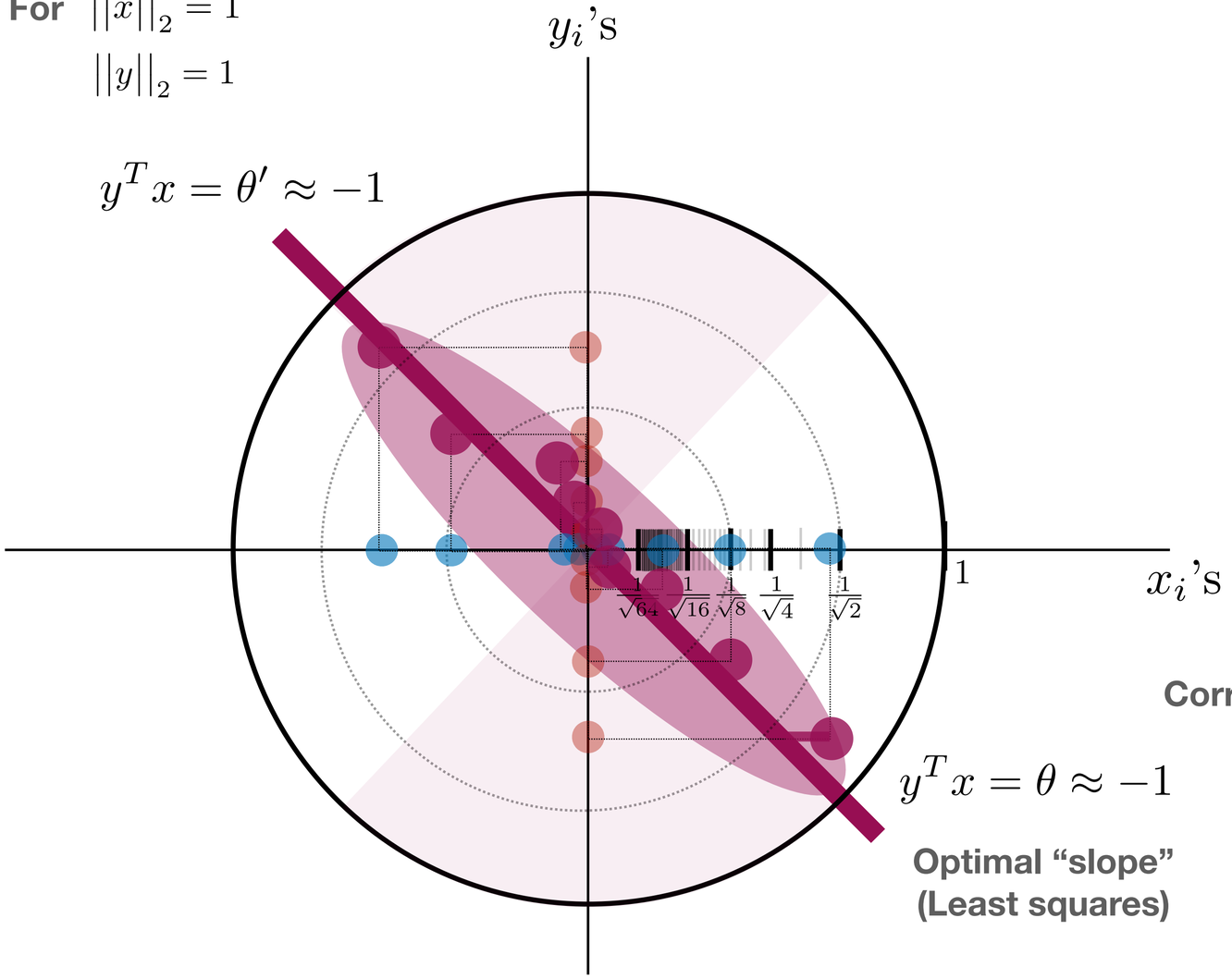
$$\min_\theta \; ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y \; = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} \; ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x \; = \frac{||x||_2}{||y||_2} \cos\theta$$

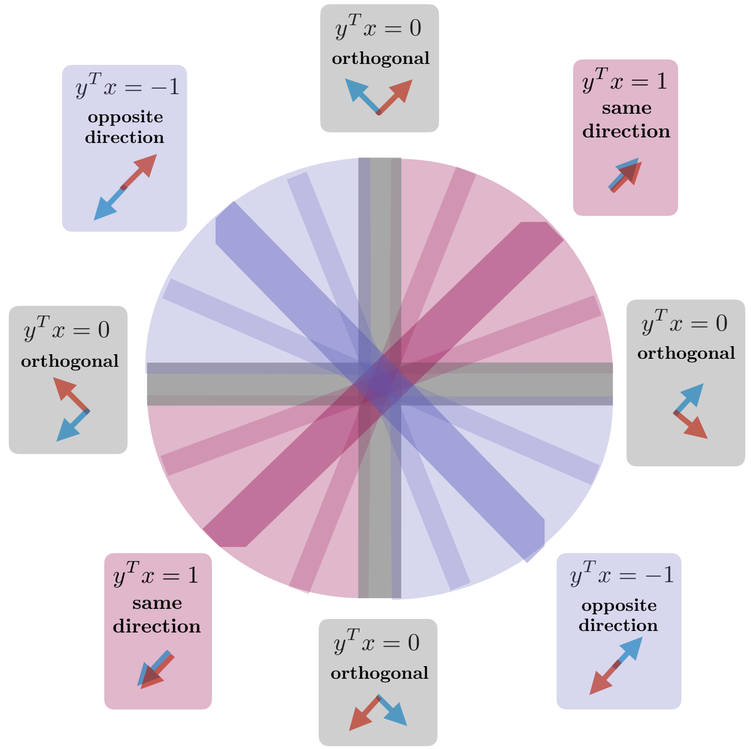
Relationship to Regression

For $||x||_2 = 1$

$||y||_2 = 1$



Optimal "slope" (Least squares)



Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

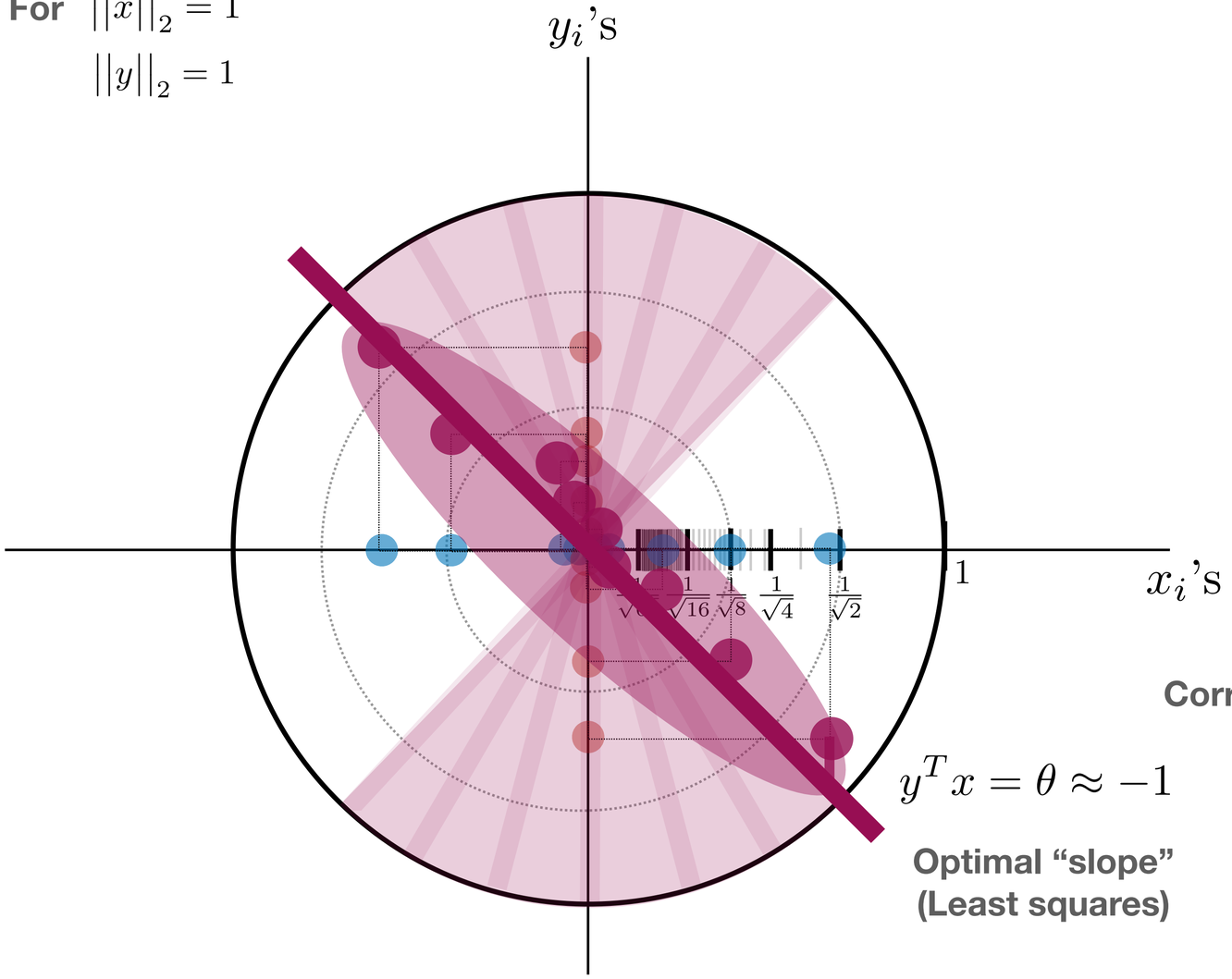
$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$

Relationship to Regression

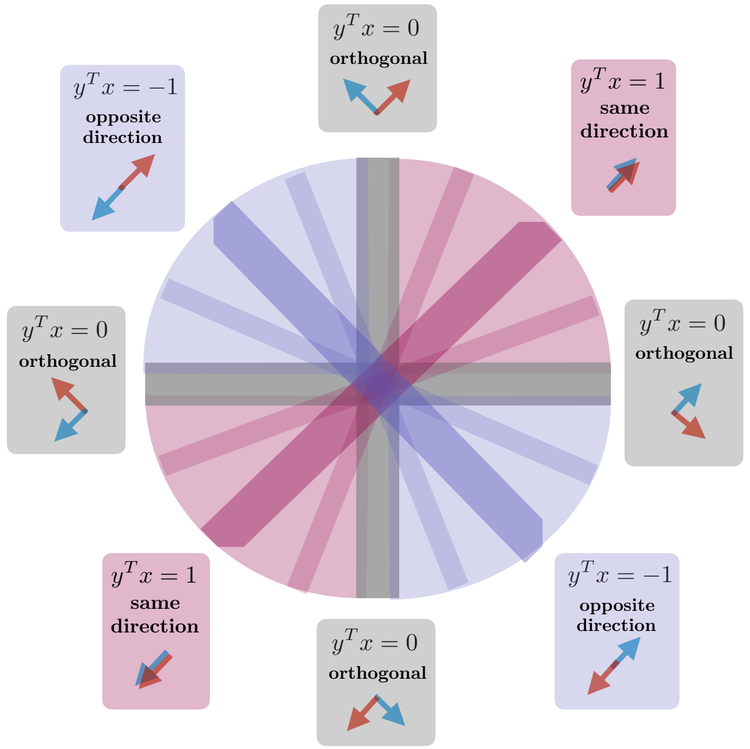
For $||x||_2 = 1$

$||y||_2 = 1$



$y^T x = \theta \approx -1$

Optimal "slope" (Least squares)



Correlation $y^T x = ||y||_2 ||x||_2 \cos\theta$

$$\min_{\theta} ||y - x\theta|| \qquad \theta = (x^T x)^{-1} x^T y = \frac{||y||_2}{||x||_2} \cos\theta$$

$$\min_{\theta'} ||x - y\theta'|| \qquad \theta' = (y^T y)^{-1} y^T x = \frac{||x||_2}{||y||_2} \cos\theta$$