Suppose A is invertible, why is $B = A^T A$ also?

- Since A is invertible so is A^T.
- The product of invertible matrices is also invertible, so B is invertible...

What's the inverse?

$$B^{-1} = (A^T A)^{-1} = A^{-1}(A^T)^{-1} = \boxed{A^{-1}(A^{-1})^T}$$

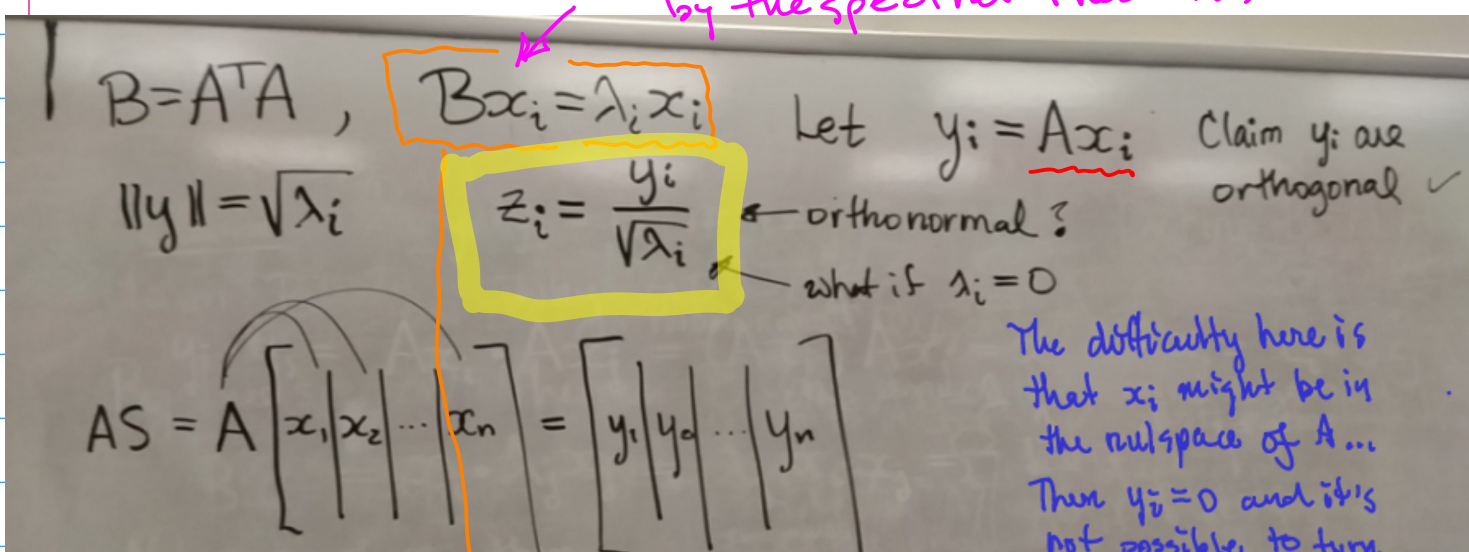
Check it

$$B B^{-1} = A^T \cancel{A}\, \cancel{A^{-1}}(A^{-1})^T = A^T (A^{-1})^T = (A^{-1} A)^T$$

$$= (I)^T = I$$

Simplifying assumption: A is invertible. That means B invertible which further means all the eigenvalues λ_i of B are non-zero.

x's are an orthonormal eigenbasis by the spectral theorem.



The difficulty here is that x_i might be in the nullspace of A... Then $y_i = 0$ and it's not possible to turn

$$y_i \cdot y_j = Ax_i \cdot Ax_j = (Ax_i)^T A x_j = x_i^T \overset{B}{A^T A} x_j$$

$$= x_i^T B x_j = x_i^T \lambda_j x_j = \lambda_j \, x_i \cdot x_j = \begin{cases} \lambda_j & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Thus $\|y_i\| = \sqrt{\lambda_i}$ and by assumption $\lambda_i \neq 0$ since A was invertible...

Define $z_i = \dfrac{y_i}{\sqrt{\lambda_i}}$ and these are orthonormal..

$$AS = A\left[\, x_1 \,\middle|\, x_2 \,\middle|\, \cdots \,\middle|\, x_n \,\right] = \left[\, y_1 \,\middle|\, y_2 \,\middle|\, \cdots \,\middle|\, y_n \,\right]$$ where $S = \left[\, x_1 \,\middle|\, x_2 \,\middle|\, \cdots \,\middle|\, x_n \,\right]$

Define $V = \left[\, z_1 \,\middle|\, z_2 \,\middle|\, \cdots \,\middle|\, z_n \,\right]$

- V is square
- V has orthonormal cols...

Thus $V^T = V^{-1}$

Recall also that $S^T = S^{-1}$.

$$V \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ 0 & & \ddots \;\; \sqrt{\lambda_n} \end{bmatrix} = \left[\, z_1 \,\middle|\, z_2 \,\middle|\, \cdots \,\middle|\, z_n \,\right] \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ 0 & & \ddots \;\; \sqrt{\lambda_n} \end{bmatrix}$$

$$= \left[\, \frac{y_1}{\sqrt{\lambda_1}} \,\middle|\, \frac{y_2}{\sqrt{\lambda_2}} \,\middle|\, \cdots \,\middle|\, \frac{y_n}{\sqrt{\lambda_n}} \,\right] \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ 0 & & \ddots \;\; \sqrt{\lambda_n} \end{bmatrix} = \left[\, y_1 \,\middle|\, y_2 \,\middle|\, \cdots \,\middle|\, y_n \,\right]$$

The same so...

Therefore,

$$AS = V\Sigma \quad \text{where } \Sigma = \begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_n} \end{bmatrix}$$

Or

$$A = V\Sigma S^T$$

← singular value decomposition of the matrix A.

usually the matrix S is called $U = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$

usually $\sigma_i = \sqrt{\lambda_i}$ so $\Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ 0 & & \ddots & \\ & & & \sigma_n \end{bmatrix}$

singular values

$$A = V\Sigma U^T$$

orthogonal matrix

diagonal.

orthogonal matrix

Treat the case where A is not invertible now and work a numerical example on Friday...

So if A is not invertible than some of the z's don't exist... so instead on n vectors, there are only m of them where $m < n$.

$$z_i = \frac{y_i}{\sqrt{\lambda_i}}$$

If $\lambda_i = 0$ can't divide here and anyway that would mean $y_i = 0$ in the first place

Assume $\lambda_i \neq 0$ for $i = 1, \ldots, m$

and $\lambda_i = 0$ for $i = m+1, \ldots, n$.

Thus

$$V = \left[z_1 \,\middle|\, z_2 \,\middle|\, \cdots \,\middle|\, z_m \,\middle|\, \text{missing vectors} \right]$$

Idea, make up some vectors and plop them in here.

add vectors that are orthonormal to the ones we already have...

$$V = \left[z_1 \,\middle|\, z_2 \,\middle|\, \cdots \,\middle|\, z_m \,\middle|\, v_{m+1} \,\middle|\, \cdots \,\middle|\, v_n \right]$$

- square
- orthonormal columns

So $V^T = V^{-1}$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_m & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

- It doesn't matter what the new vectors are because they are multiplied by the zeros in Σ anyway and disappear.

$$A = V \Sigma U^{-1}$$