# Medical Image Analysis and Processing

## Image Noise Filtering

## Sparse Denoising

Emad Fatemizadeh

Distance/online Course: Session 14 (Episode #2)

Date: 13 April 2021, 24th Farvardin 1400

# Contents

# Norm Definition

- Consider $x \in \mathbb{R}^n$
- $p$-norm:

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}, \qquad p > 0$$

- Euclidean Norm, 2-Norm, $\ell_2$ Norm:

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

- Manhattan Norm, 1-Norm, $\ell_1$ Norm

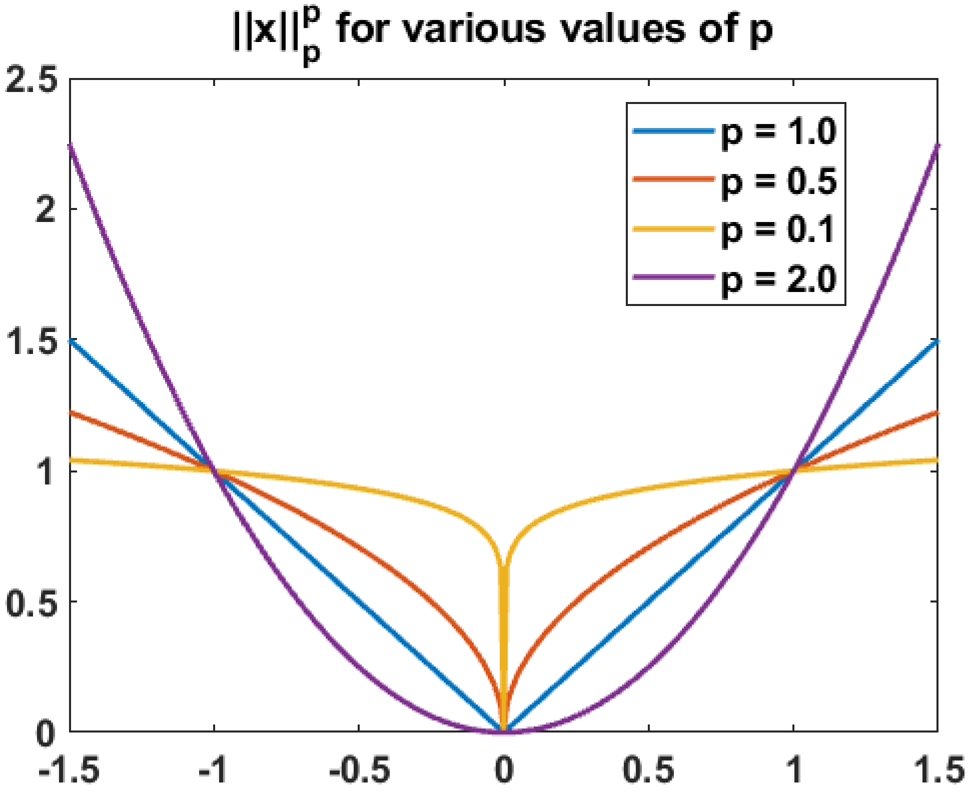
$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

- zero-Norm, $\ell_0$ Norm (number of non-zero entries of the vector $x$):

$$\|x\|_0 = |x_1|^0 + |x_2|^0 + \cdots + |x_n|^0, \qquad 0^0 \stackrel{\text{def}}{=} 0$$

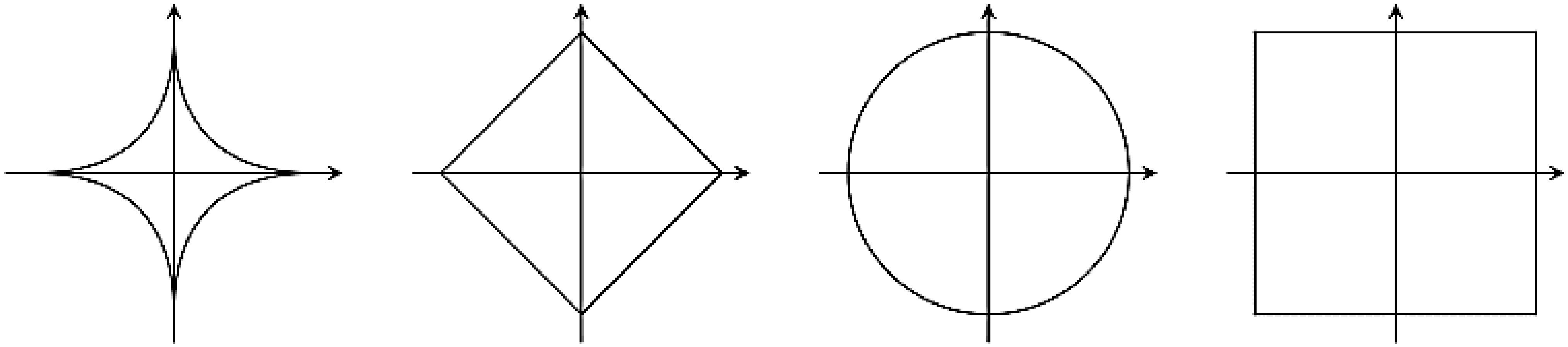
# Norm Definition

› The behavior of $||x||_p^p$ for various values of $p$



# Norm Definition

› Unit Circles ($\|x\|_p$=1) in different norms:



$p = \frac{1}{2}$ $p = 1$ $p = 2$ $p = \infty$

# Sparse Approximation

› Start with an under-determined set of equations:

$$D_{n\times k} x_{k\times 1} = y_{n\times 1}, \qquad k > n$$

› $n$ -equations and $k$ -unknown → infinitely many solutions

› Example:

$$\begin{cases} x_1 + 3x_2 - x_3 = y_1 \\ x_1 - 8x_2 + 2x_3 = y_2 \end{cases}, n = 2, k = 3$$

$$\begin{bmatrix} 1 & 3 & -1 \\ 1 & -8 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

# Sparse Approximation

› Signal processing perspective of set-of- equations

$$D_{n\times k}x_{k\times 1} = y_{n\times 1}, \qquad k > n$$

› $y_{n\times 1}$: signal

› $x_{k\times 1}$: representation of signal

› $D_{n\times k}$: Dictionary

› An old fashion example (but determined, $k = n$):

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-jk\frac{2\pi}{N}} \Leftrightarrow x(n) = \sum_{k=0}^{N-1} X(k)e^{jk\frac{2\pi}{N}}$$

# Sparse Approximation

› An old fashion example (but determined):

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-jk\frac{2\pi}{N}} \Leftrightarrow x(n) = \sum_{k=0}^{N-1} X(k) e^{jk\frac{2\pi}{N}}$$

› In matrix formulation:

$$\boldsymbol{X} = \boldsymbol{W}^H \boldsymbol{x} \Leftrightarrow \boldsymbol{x} = \boldsymbol{W}\boldsymbol{X}$$

› Where:

- $\boldsymbol{X} = [X(0) \quad X(1) \quad \dots \quad X(N-1)]^T$
- $\boldsymbol{x} = [x(0) \quad x(1) \quad \dots \quad x(N-1)]^T$
- $\boldsymbol{W} = \left[e^{jk\frac{2\pi}{N}n}\right]_{k,n=0}^{N-1}$

# Sparse Approximation

› What is dictionary role?

$$D_{n\times k}x_{k\times 1} = y_{n\times 1}$$

› Let show D with its columns (atoms):

$$D = [\boldsymbol{d}_1|\boldsymbol{d}_2|\dots|\boldsymbol{d}_k]$$

$$y = Dx = [\boldsymbol{d}_1|\boldsymbol{d}_2|\dots|\boldsymbol{d}_k]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \sum_{i=1}^{k} x_i \boldsymbol{d}_i$$

› Signal ($y$) represented by an expansion using dictionary atoms ($\boldsymbol{d}_i$'s) as basis and ($x_i$'s) as coefficients

# Sparse Approximation

- Example

$$\begin{bmatrix} 1 & 3 & -1 \\ 1 & -8 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 3 \\ -8 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_3$$

- Thus we may use ($x_i$’s) instead of ($y_i$’s)

# Sparse Approximation

- Let find solution(s) for:

$$D_{n\times k}x_{k\times 1} = y_{n\times 1}, \qquad k > n$$

- Obviously there is infinitely many solutions.
- We solve the following constrained optimization problems:

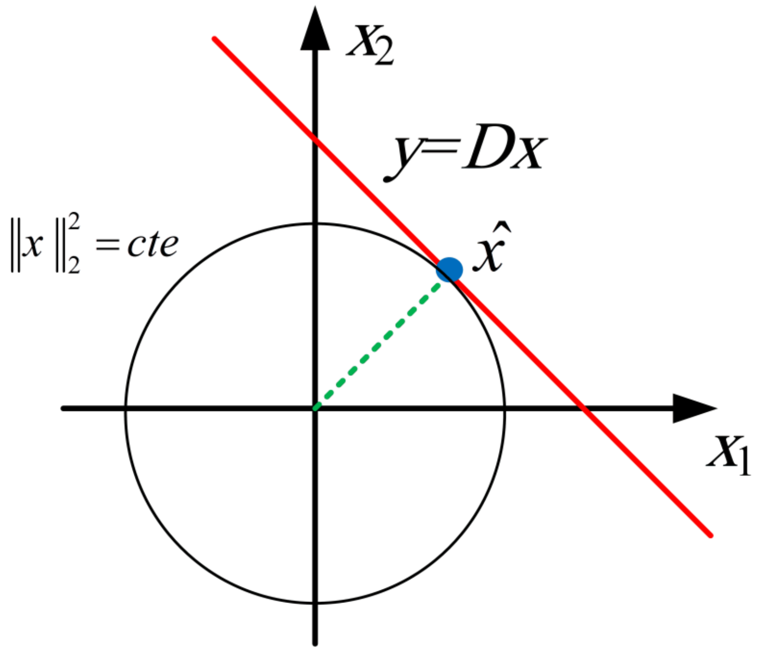
$$\hat{x} = \underset{x}{\operatorname{argmin}}\, J(x), \qquad s.t.\ Dx = y$$

- Cost function, $J$, depends on problems

# Sparse Approximation

- Example:

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|x\|_2^2, \qquad s.t.\ Dx = y \Rightarrow \hat{x} = D^{+}y = D^{T}(DD^{T})^{-1}y$$

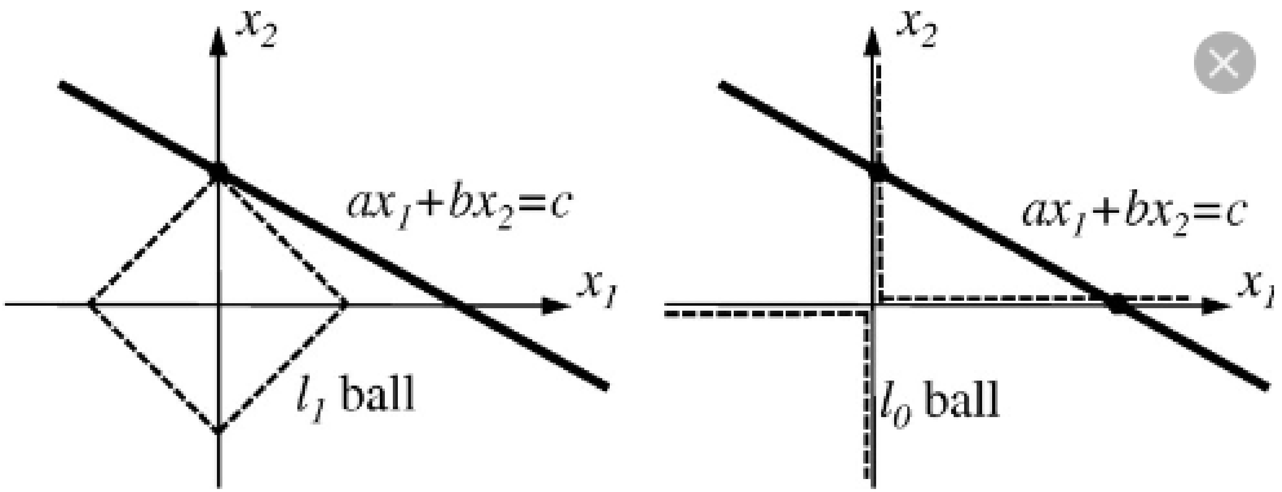


# Sparse Approximation

› Other solution(s):

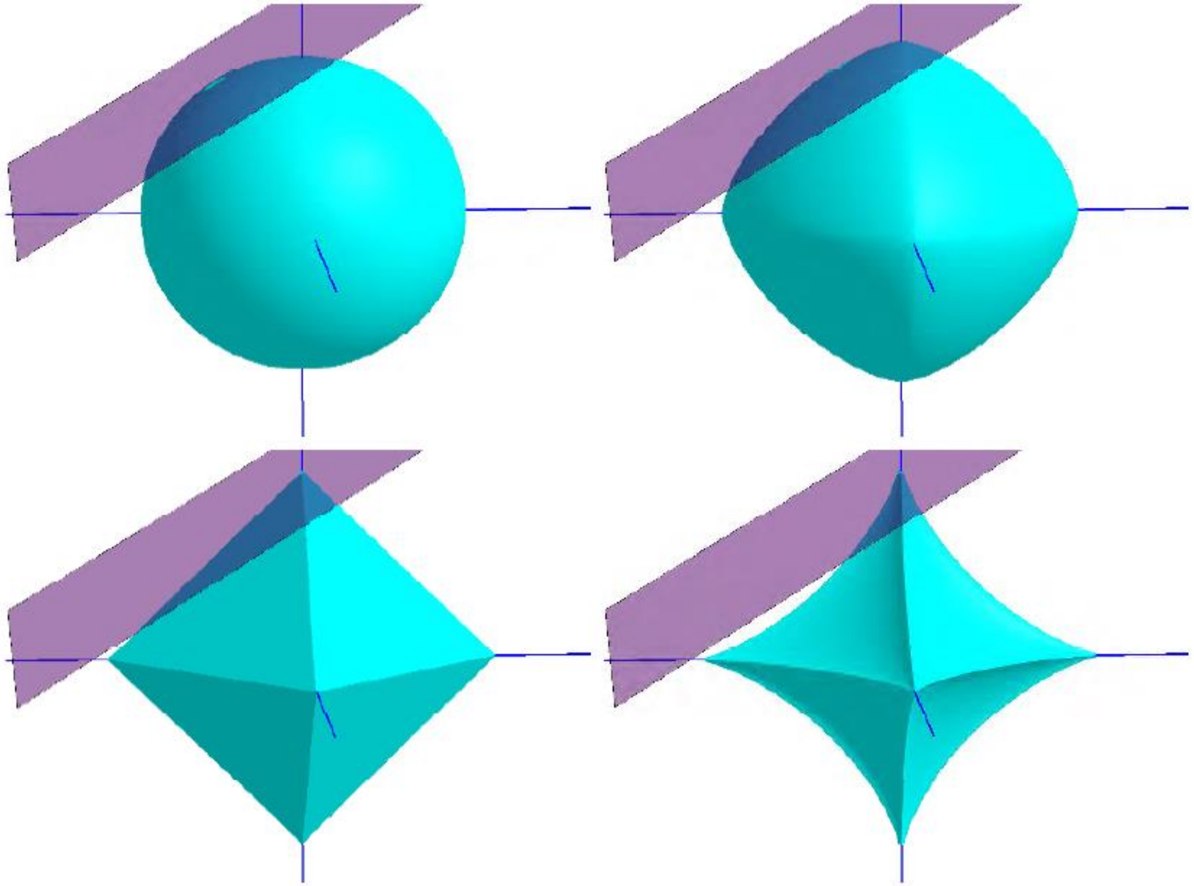
$$\hat{x} = \underset{x}{\operatorname{argmin}} \|x\|_1, \qquad s.t.\ Dx = y$$

$$\hat{x} = \underset{x}{\operatorname{argmin}} \|x\|_0, \qquad s.t.\ Dx = y$$



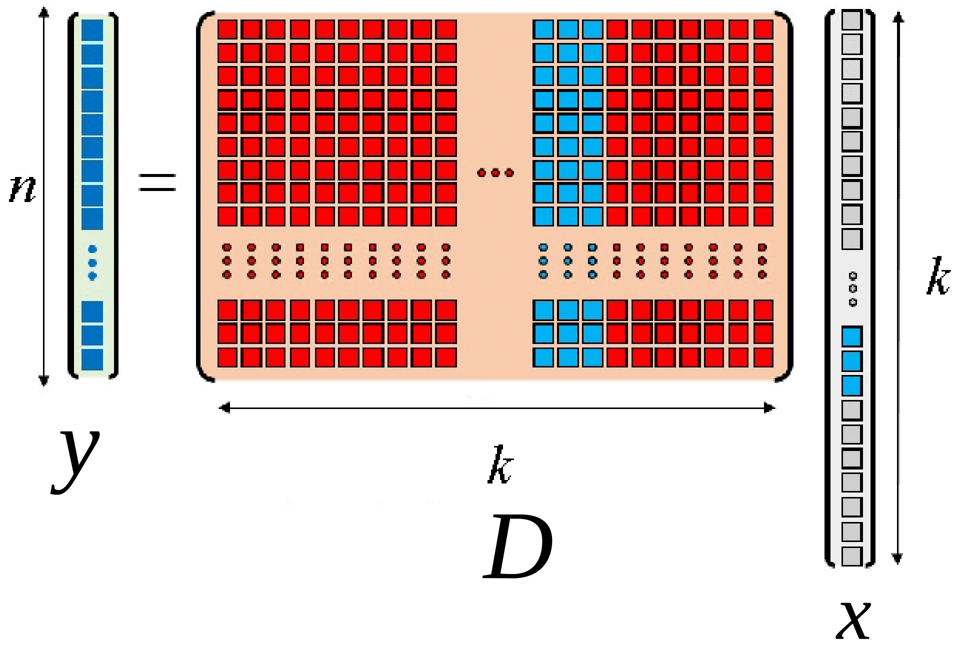
# Sparse Approximation

› 3D illustration:



# Sparse Approximation

› We focus on $\ell_0$-norm minimization, why? Sparsity!




› High dimension representation ($x$) may be sparse for a well-chosen dictionary, D

# Sparse Approximation

- How to solve:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\{\|x\|_0\}, \qquad s.t.\ Dx = y$$

- $|x|_0$ is too sensitive to noise, perfect solution is impossible, we solve an alternative:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\{\|x\|_0\}, \qquad s.t. \|Dx - y\|_2^2 \leq \varepsilon$$

- where $\varepsilon$ is closely related to the properties of the noise.
- Or alternatively:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\{\|Dx - y\|_2^2\}, \qquad s.t. \|x\|_0 \leq T_0$$

# Sparse Approximation

› Several approach for solving $\ell_0$-norm has been proposed:

› Relaxation approach are well-known (that smooth the $\ell_0$-norm and solve the alternative problem):

› Method #1) convexise" the $\ell_0$-norm and replace it with $\ell_1$-norm

$$\hat{x} = \underset{x}{\operatorname{argmin}}\{\|x\|_1\}, \qquad s.t.\ \|Dx - y\|_2^2 \leq \varepsilon$$

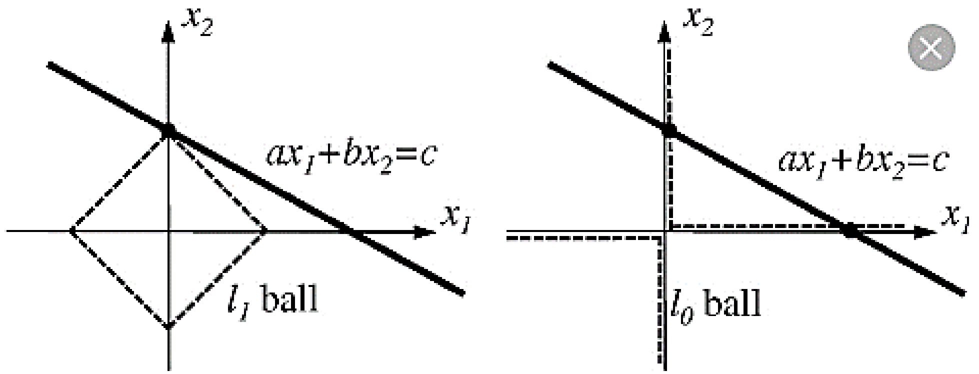
› For an appropriate Lagrange multiplier $\lambda$, we may solve this equivalent problem:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\left\{\frac{1}{2}\|Dx - y\|_2^2 + \lambda\|x\|_1\right\}$$

› Where $\lambda$ is function of $D$, y, and $\varepsilon$

# Sparse Approximation

› Why this may works:




› Method #2) replace $\|x\|_0$ with:

$$\sum_{i=1}^{k}\left(1-exp\left(-\beta x_i^2\right)\right), \beta \to \infty$$

# Dictionary

- Dictionary is too important to  generate sparse representation
- Example:
- Sparse representation of a signal with just one non-zero value:

$$y = \delta[n - n_0], y \in \mathbb{R}^N \Rightarrow y = I_{N\times N}x, \qquad x \in \mathbb{R}^N$$

- where $I_{N\times N}$ is identity matrix (Dirac dictionary) of size N
- Sparse representation of a signal with just two non-zero values:

$$y = sin[\omega_0 n], y \in \mathbb{R}^N \Rightarrow y = W_{N\times N}x, \qquad x \in \mathbb{R}^N$$

- where$W_{N\times N}$is inverse Fourier matrix of size N

# Dictionary

- What about combination of two signals:

$$y = sin[\omega_0 n] + \delta[n - n_0]?$$

- Solution is an overcomplete dictionary and under-determined representation:

$$y = \underbrace{[I_{N\times N}|W_{N\times N}]}_{D} x, x \in \mathbb{R}^{2N}, D \in \mathbb{R}^{N\times 2N}$$

- How to determine dictionary:
  - Building
  - Learning

# Dictionary

- Example:

$$\begin{bmatrix} 0.6533 \\ 0.2706 \\ -0.2706 \\ -0.6533 \end{bmatrix} = \begin{bmatrix} 0.5000 & 0.6533 & 0.5000 & 0.2706 \\ 0.5000 & 0.2706 & -0.5000 & -0.6533 \\ 0.5000 & -0.2706 & -0.5000 & 0.6533 \\ 0.5000 & -0.6533 & 0.5000 & -0.2706 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

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

# Dictionary

› Example, Now consider:

$$\begin{bmatrix} 0.6533 \\ 0.2706 \\ -0.2706 \\ -0.6533 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5000 & 0.6533 & 0.5000 & 0.2706 \\ 0.5000 & 0.2706 & -0.5000 & -0.6533 \\ 0.5000 & -0.2706 & -0.5000 & 0.6533 \\ 0.5000 & -0.6533 & 0.5000 & -0.2706 \end{bmatrix} \underbrace{\begin{bmatrix} 0.5000 \\ 0.7294 \\ -0.5000 \\ 0.6533 \end{bmatrix}}_{\|x\|_0=4}$$

$$\begin{bmatrix} 0.6533 \\ 0.2706 \\ -0.2706 \\ -0.6533 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 0.6533 \\ 0.2706 \\ 0.7294 \\ -0.6533 \end{bmatrix}}_{\|x\|_0=4}$$

# Dictionary

› Example, But:

$$\begin{bmatrix}0.6533\\0.2706\\-0.2706\\-0.6533\end{bmatrix}+\begin{bmatrix}0\\0\\1\\0\end{bmatrix}=\begin{bmatrix}0.5000 & 0.6533 & 0.5000 & 0.2706 & 1 & 0 & 0 & 0\\0.5000 & 0.2706 & -0.5000 & -0.6533 & 0 & 1 & 0 & 0\\0.5000 & -0.2706 & -0.5000 & 0.6533 & 0 & 0 & 1 & 0\\0.5000 & -0.6533 & 0.5000 & -0.2706 & 0 & 0 & 0 & 1\end{bmatrix}\underbrace{\begin{bmatrix}0\\\mathbf{1}\\0\\0\\0\\\mathbf{1}\\0\\0\end{bmatrix}}_{\|x\|_0=2}$$

# The End

- AnY QuEsTiOn?