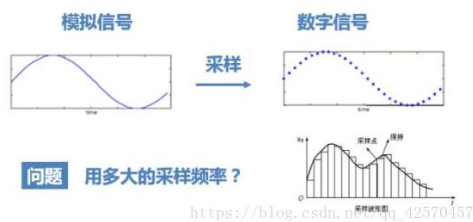


一、什么是压缩感知（CS）？

compressed sensing 又称 compressed sampling, CS 是一个针对信号采样的技术, 它通过一些手段, 实现了“压缩的采样”, 准确说是在采样过程中完成了数据压缩的过程。

因此我们首先要从信号采样讲起:

模拟信号采样



1. 我们知道, 将模拟信号转换为计算机能够处理的数字信号, 必然要经过采样的过程。问题在于, 应该用多大的采样频率, 即采样点应该多密多疏, 才能完整保留原始信号中的信息呢?

奈奎斯特采样定理



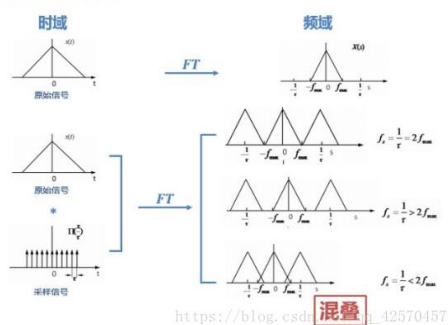
奈奎斯特, N.
1889-1976

要想让采样之后的数字信号完整保留原始信号中的信息, 采样频率必须大于信号中最高频率的2倍!

https://blog.csdn.net/qq_42570457

2. 奈奎斯特给出了答案——信号最高频率的两倍。一直以来, 奈奎斯特采样定律被视为数字信号处理领域的金科玉律。

为什么是这样？



3. 至于为什么是两倍，学过信号处理的同学应该都知道，时域以 τ 为间隔进行采样，频域会以 $1/\tau$ 为周期发生周期延拓。那么如果采样频率低于两倍的信号最高频率，信号在频域频谱搬移后就会发生**混叠**。

一定要如此吗？

2004年，几位大牛证明，如果信号是**稀疏的**，那么它可以由**远低于**采样定理要求的采样点重建恢复，并于**2007年**正式提出了“**压缩感知**”（Compressed Sensing）这个概念。

压缩感知



陶哲轩



Emmanuel Candes



David Donoho

4. 然而这看似不容置疑的定律却受到了几位大神的挑战。Candes 最早意识到了突破的可能，并在不世出的数学天才陶哲轩以及 Candes 的老师 Donoho 的协助下，提出了压缩感知理论，该理论认为：**如果信号是稀疏的，那么它可以由远低于采样定理要求的采样点重建恢复。**

压缩感知

“采样频率须大于信号中最高频率的2倍”——Nyquist

采样频率 $\xrightarrow{\text{意味着}}$ 等间距采样

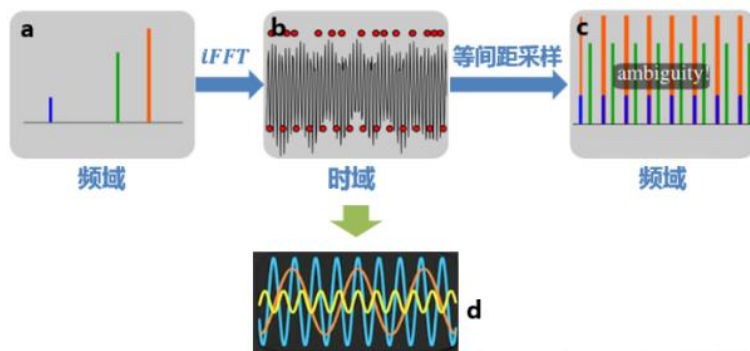
- 等间距采样，频域将以 $1/\tau$ 为周期延拓，采样频率低必然引起混叠。
- 如果是不等间距采样呢？
- 如果是随机采样呢？
https://blog.csdn.net/qq_42570457

5. 而突破的关键就在于采样的方式。当我们说“采样频率”的时候，意味着做的是等间距采样，数字信号领域通常都是做等间距采样，也服从奈奎斯特采样定律。

但是如果是不等间距采样呢？依然必须要服从采样定理吗？

压缩感知

采样



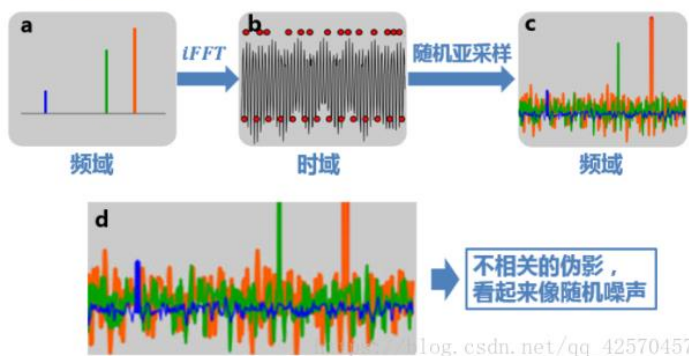
https://blog.csdn.net/qq_42570457

6. 答案是，随机的亚采样给了我们恢复原信号的可能。

上图非常关键，它可以简单直观地表述压缩感知的思路。如图 b、d 为三个余弦函数信号叠加构成的信号，在频域的分布只有三条线（图 a）。如果对其进行 8 倍于全采样的等间距亚采样（图 b 下方的红点），则频域信号周期延拓后，就会发生混叠（图 c），无法从结果中复原出原信号。

压缩感知

采样



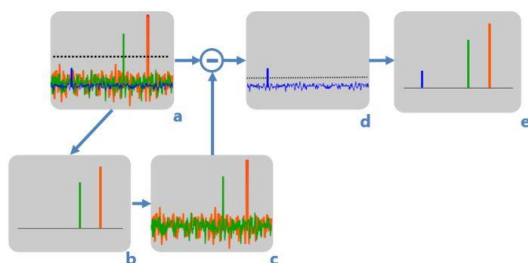
7. 而如果采用**随机亚采样**（图 b 上方的红点），那么这时候频域就不再是以固定周期进行延拓了，而是会产生大量不相关（incoherent）的干扰值。如图 c，最大的几个峰值还依稀可见，只是一定程度上被干扰值覆盖。这些干扰值看上去非常像随机噪声，但实际上是由于三个原始信号的非零值发生能量泄露导致的（不同颜色的干扰值表示它们分别是由于对应颜色的原始信号的非零值泄露导致的）

P.S：为什么随机亚采样会有这样的效果？

这可以理解成随机采样使得频谱不再是整齐地搬移，而是一小部分一小部分胡乱地搬移，频率泄露均匀地分布在整個频域，因而泄露值都比较小，从而有了恢复的可能。

压缩感知

恢复



8. 接下来的关键在于，信号该如何恢复？下面讲一种典型的算法（匹配追踪）：

(1) 由于原信号的频率非零值在亚采样后的频域中依然保留较大的值，其中较大的两个可以通过设置阈值，检测出来（图 a）。

(2) 然后，假设信号只存在这两个非零值（图 b），则可以计算出由这两个非零值引起的干扰（图 c）。

(3) 用 a 减去 c，即可得到仅由蓝色非零值和由它导致的干扰值(图 d)，再设置阈值即可检测出它，得到最终复原频域（图 e）

(4) 如果原信号频域中有更多的非零值，则可通过迭代将其一一解出。

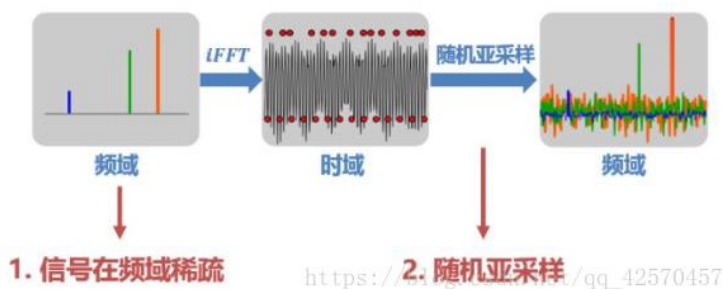
以上就是压缩感知理论的核心思想——以比奈奎斯特采样频率要求的采样密度更稀疏的密度对信号进行随机亚采样，由于频谱是均匀泄露的，而不是整体延拓的，因此可以通过特别的追踪方法将原信号恢复。

二、压缩感知的前提条件

接下来我们总结一下，能实现压缩感知的关键在于什么，即需要哪些前提条件。

前提条件

刚才的例子中，满足了两个前提条件：



9. 在刚才的讲述中大家可以感受到，这个例子之所以能够实现最终信号的恢复，是因为它满足了两个前提条件：

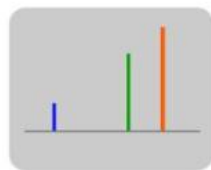
(1). 这个信号在频域只有 3 个非零值，所以可以较轻松地恢复出它们。

(2). 采用了随机亚采样机制，因而使频率泄露均匀地分布在整個频域。

这两点对应了 CS 的两个前提条件——稀疏性 (sparsity)、不相关性 (incoherence)。

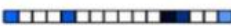
前提条件-1 稀疏性

信号需要在某一个变换域具有**稀疏性**



信号在频域是**稀疏的**

如果信号在某个域中**非零点远远小于信号总点数**，
则信号在该域中是**稀疏的**。

x 

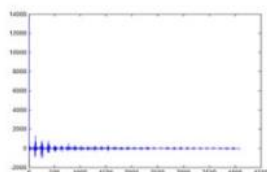
典型的一维稀疏信号，只有少量非零项

10. 关于**稀疏性**可以这样简单直观地理解：若信号在某个域中**只有少量非零值**，那么它在该域稀疏，该域也被称为信号的**稀疏域**。

因此，第一个前提条件要求信号必须在某一个变换域具有稀疏性。比如例子中，信号在频域是稀疏的，因而可以通过所述的重建方法轻松地在稀疏域（频域）复原出原信号。

前提条件-1 稀疏性

对于压缩感知：



信号只需要近似满足稀疏性，即为可压缩信号。

对于CS，只要它在某一个变换域满足近似稀疏特性即可，我们称之为稀疏域，重建将在稀疏域进行。



Ψ 可以是频域、小波变换、离散余弦变换等

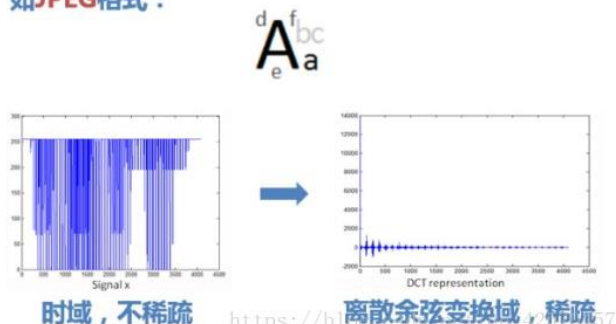
然而通常信号在变换域中不会呈现完全的稀疏性。其实只要它近似满足稀疏性，即大部分值趋于零，只有少量大的非零值，就可以认为它是可压缩信号，可以对它进行CS亚采样。

对于之前讲的例子，如果它在频域中不稀疏，我们可以做DWT、DCT等，找到它的稀疏变换。

补充知识 稀疏性

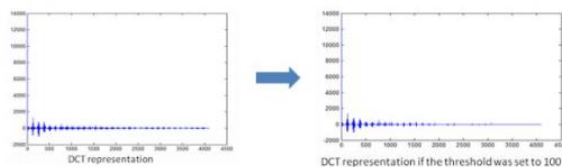
信号的稀疏性已经在图像压缩领域有了很好的应用。

如JPEG格式：



11. 这里针对信号的稀疏性和信号压缩额外补充一下：其实，信号的稀疏性已经在图像压缩领域有了很广泛的应用。利用信号的稀疏性，可以对信号进行压缩。如图像压缩领域的JPEG格式，就是将图像变换到离散余弦域，得到近似稀疏矩阵，只保留较大的值，从而实现压缩。

补充知识 稀疏性



舍弃小于100的值

$A_{e,a}^{d,f,bc}$ → $A_{e,a}^{d,f,bc}$ 仅占原图像数据大小的6.9%

4096个点 282个点

https://blog.csdn.net/qq_42570457

12. 比如这个例子中，仅用原图像 6.9% 的点就复原了和原图像基本相同的图像。我们还可以采用小波变换，即为 JPEG-2000，压缩效果更好。

补充知识 图像压缩与压缩感知

图像压缩



压缩感知的原动力问题：

既然全采样了还要再丢弃，我们为什么不能直接少采样一些点？

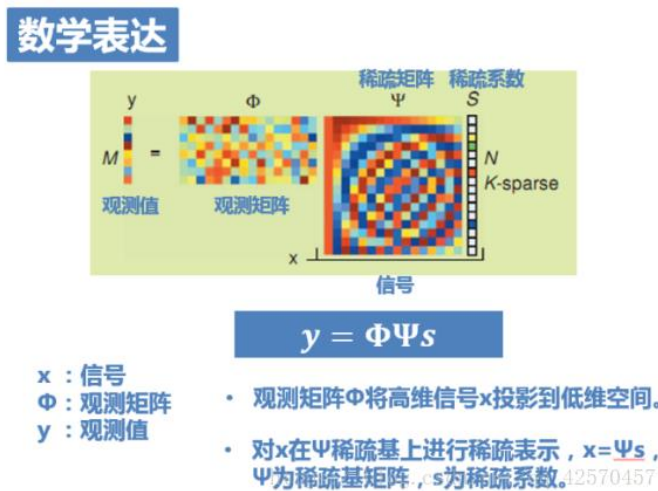
压缩感知



13. 这里需要指出，图像压缩和压缩感知这两个概念很容易弄混，大家一定要分清。

它们其实有着本质上的区别。图像压缩是先进行了全采样，然后再变换域丢弃小系数，完成压缩；

而压缩感知不同，它的思想其实从图像压缩中借鉴了很多：既然全采样了还要再丢弃，我们为什么不能直接少采样一些点？因此，压缩感知直接进行了亚采样，然后再用算法消除亚采样导致的伪影。可以说，压缩感知**直接在采样时就完成了压缩**。



14. 接下来，在将第二个前提条件之前，还是需要引入必要的数学表达的。上图是一个大家在压缩感知相关的书籍文献中会经常看到的一张示意图。很多文章试图用这张图给大家讲清楚什么是压缩感知，结果导致大家看得一头雾水，混淆在各种“矩阵”当中。不过相信有了我之前的讲解，现在这张图会好理解很多。这张图也就是把亚采样的过程用矩阵的方式表达出来而已：

如图， x 是为长度 N 的一维信号，也就是原信号，稀疏度为 k 。此刻它是未知的。

Φ 为观测矩阵，对应着亚采样这一过程。它将高维信号 x 投影到低维空间，是已知的。

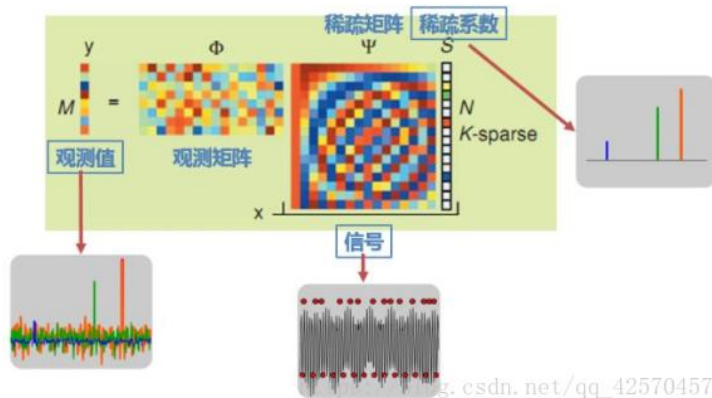
$y = \Phi x$ 为长度 M 的一维测量值，也就是亚采样后的结果。显然它也是已知的。

因此，压缩感知问题就是在已知测量值 y 和测量矩阵 Φ 的基础上，求解欠定方程组 $y = \Phi x$ 得到原信号 x 。

然而，一般的自然信号 x 本身并不是稀疏的，需要在某种稀疏基上进行稀疏表示。令 $x = \Psi s$ ， Ψ 为稀疏基矩阵， s 为稀疏系数。

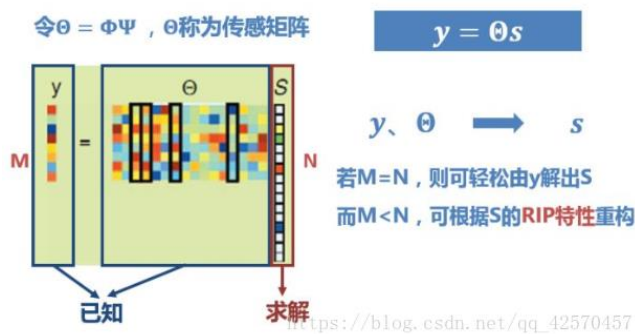
于是最终方程就变成了： $y = \Phi \Psi s$ 。已知 y 、 Φ 、 Ψ ，求解 s 。

数学表达



15. 对应一开始的例子大家就能明白： x 就是三个正弦信号叠加在一起的原信号；稀疏矩阵 Ψ 就是傅里叶变换，将信号变换到频域 S ；而观测矩阵 Φ 就对应了我们采用的随机亚采样方式； y 就是最终的采样结果。

数学表达



16. $y = \Phi \Psi s$ 有点长，我们把 $\Phi \Psi$ 合并成一个矩阵，称之为传感矩阵。即令 $\Theta = \Phi \Psi$ ，则 $y = \Theta s$ 。

问题即为，已知 y 和 Θ ，求解 s 。

求解出 S 后，由 $x = \Psi S$ 即可得到恢复出的原信号 x 。

然而在正常情况下，方程的个数远小于未知数的个数，方程是没有确定解的，无法重构信号。但是，由于信号是 K 稀疏，如果上式中的 Φ 满足有限等距性质 (RIP)，则 K 个系数就能够从 M 个测量值准确重构（得到一个最优解）。

前提条件-2 不相关

之前已提到，采用随机采样才能实现信号的恢复。



陶哲轩和Candès于2005年给出了更为准确的要求：观测矩阵 Φ 应满足约束等距性条件 (Restricted Isometry Property, 简称RIP)：



即对于任意和常数，有：

$$(1 - \delta_k) \|c\|_2^2 \leq \|\phi c\|_2^2 \leq (1 + \delta_k) \|c\|_2^2$$

https://blog.csdn.net/qq_42570457

Baraniuk证明：

RIP的等价条件是观测矩阵和稀疏表示基不相关 (incoherent)



$$y = \Phi \Psi s$$

$$\Phi \xleftrightarrow{\text{不相关}} \Psi$$

相关性的定义：

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \leq k, j \leq n} |\langle \phi_k, \psi_j \rangle|.$$

μ 的范围： $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$

μ 越小， Φ 和 Ψ 越不相关

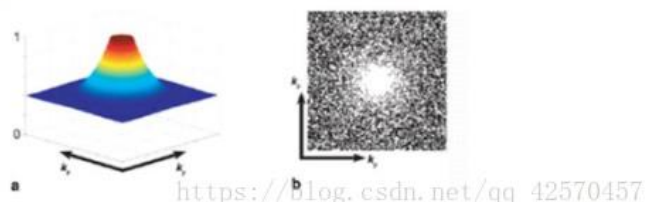
https://blog.csdn.net/qq_42570457

17. 接下来的数学内容可以简短略过：陶大神和 Candès 大神证明了 RIP 才是观测矩阵要满足的准确要求。但是，要确认一个矩阵是否满足 RIP 非常复杂。于是 Baraniuk 证明：RIP 的等价条件是观测矩阵和稀疏表示基不相关 (incoherent)。

这就是压缩感知的第二个前提条件。

陶哲轩和Candès证明：

独立同分布的高斯随机测量矩阵可以成为普适的压缩感知测量矩阵。



18. 那怎样找到不相关的观测矩阵呢？陶哲轩和 Candès 又证明：独立同分布的高斯随机测量矩阵可以成为普适的压缩感知测量矩阵。

于是满足高斯分布的随机测量矩阵就成了 CS 最常用的观测矩阵。

对于二维信号，往往就采用如右上图所示的采样矩阵对图像进行亚采样。

对于一维信号，采用前文提到的随机不等间距的亚采样即可。

到这里，我们可以这样用一句话概括地描述什么是压缩感知：

如果一个信号在某个变换域是稀疏的，那么就可以用一个与变换基不相关的观测矩阵将变换所得高维信号投影到一个低维空间上，然后通过求解一个优化问题就可以从这些少量的投影中以高概率重构出原信号。

以上可以算作是压缩感知的定义吧。但是如果要再简洁一点呢？

在我看来，压缩感知可以用这样一句话来表述：

直接采集出一个 JPEG

——之前图像压缩的方法是全采样之后再压缩，抛弃稀疏变换域中的一些小系数；而 CS 直接减少了采样点，采集完后、经过重建的图像，就是一副在某变换域稀疏的压缩图像，比如 JPEG。

那这么做有什么优势呢？

对于很多情形，比如照相机拍摄照片，这样减少采样点并没有优势。因为所有像素的采集在一瞬间就都完成了。

但是对于一些采集比较慢的情形，比如核磁共振成像，CS 就可以发挥巨大优势。原本一副 MRI 图像常常需要几十秒，速度慢也是 MRI 的一大缺陷。而应用 CS 技术后，只需要采集全采样几分之一的数据，就可以重建出原图。这样就可以把成像速度提高好几倍，同时对图像质量影响不大。

另一个应用是 Rice 大学开发的单像素相机，也就是说这种相机只需要一个像素，非常有趣。感兴趣的朋友可以自己去调查。

Rice 大学关于压缩感知收集的一些论文：<http://dsp.rice.edu/cs/>

压缩感知理论主要包括三部分：

- (1) 信号的稀疏表示；
- (2) 设计测量矩阵，要在降低维数的同时保证原始信号 x 的信息损失最小；
- (3) 设计信号恢复算法，利用 M 个观测值无失真地恢复出长度为 N 的原始信号。

理论依据：

- (1) 设长度为 N 的信号 X 在某个正交基 Ψ 上是 K -稀疏的（即含有 k 个非零值）；
- (2) 如果能找到一个与 Ψ 不相关（不相干）的观测基 Φ ；
- (3) 用观测基 Φ 观测原信号得到长度 M 的一维测量值 M 个观测值 Y ， $K < M < N$ ；
- (4) 那么就可以利用最优化方法从观测值 Y 中大概率恢复 X 。



数学表达：

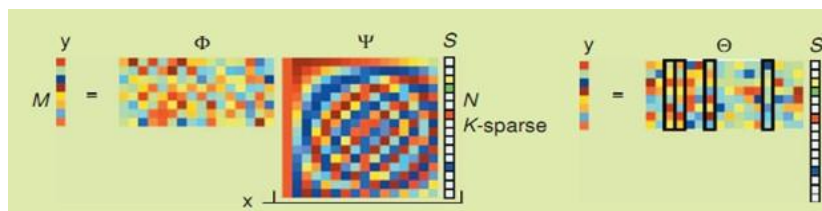
设 x 为长度 N 的一维信号，稀疏度为 k （即含有 k 个非零值）， A 为 $M \times N$ 的二维矩阵（ $M < N$ ）， $y = \Phi x$ 为长度 M 的一维测量值。压缩感知问题就是已知测量值 y 和测量矩阵 Φ 的基础上，求解欠定方程组 $y = \Phi x$ 得到原信号 x 。 Φ 的每一行可以看作是一个传感器（Sensor），它与信号相乘，拾取（Acquisition）

了信号的一部分信息。而这一部分信息足以代表原信号，并能找到一个算法来高概率恢复原信号。

一般的自然信号 x 本身并不是稀疏的，需要在某种稀疏基上进行稀疏表示， $x = \Psi s$ ， Ψ 为稀疏基矩阵， s 为稀疏系数（ s 只有 K 个是非零值（ $K \ll N$ ））。

压缩感知方程为 $y = \Phi x = \Phi \Psi s = \Theta s$ 。

将原来的测量矩阵 Φ 变换为 $\Theta = \Phi \Psi$ （称之为传感矩阵），解出 s 的逼近值 s' ，则原信号 $x' = \Psi s'$ 。



1、信号的稀疏表示

信号的稀疏性简单理解为信号中非 0 元素数目较少，或者说大多数系数为 0（或者绝对值较小）。

自然界存在的真实信号一般不是绝对稀疏的，而是在某个变换域下近似稀疏，即可压缩信号。或者说从理论上讲任何信号都具有可压缩性，只要能找到其相应的稀疏表示空间，就可以有效地进行压缩采样。信号的稀疏性或可压缩性是压缩感知的重要前提和理论基础。

稀疏表示的意义：只有信号是 K 稀疏的（且 $K \ll M \ll N$ ），才有可能在观测 M 个观测值时，从 K 个较大的系数重建原始长度为 N 的信号。也就是当信号有稀疏展开时，可以丢掉小系数而不会失真。

我们知道，长度为 N 的信号 x 可以用一组基 $\Psi^T = [\Psi_1, \dots, \Psi_M]$ 的线性组合来表示：

$x = \Psi s$ ， Ψ 为稀疏基 $N \times N$ 矩阵， s 为稀疏系数（ N 维向量），当信号 x 在某个基 Ψ 上仅有 $K \ll N$ 个非零系数或远大于零的系数 s 时，称 Ψ 为信号 x 的稀疏基。我们需要做的就是合理地选择稀疏基，使得信号的稀疏系数个数尽可能少。

再啰嗦点的话：如果长度为 N 的信号 x ，在变换域 Φ 中只有 K 个系数不为零（或者明显大于其他系数），且 $K \ll N$ ，那么可以认为信号 x 在 Φ 域中是稀疏的并可称为 K -稀疏（不是严格的定义）。那么在该域下，我们如果只保留这 M 个大系数，丢弃其他的系数，则可以减小储存该信号需要的空间，达到了压缩（有损压缩）的目的。同时，以这 M 个系数可以重构原始信号 x ，不过一般而言得到的是 x 的一个逼近。

我们应该熟悉 JPEG 跟 JPEG2000 的区别吧，JPEG 的核心算法是 DCT，而后者是 DWT，本质上，这两种处理方法都是将信号从一个域变换到另外一个域（把坐标系进行旋转，将信号投影到不同的基上），从而获得信号的稀疏表示，即用最少的系数来表示信号，不过 DWT 比 DCT 更加稀疏而已。信号不同，对应最稀疏表达的基也会不同，比如，对于一维信号可能小波基是最稀疏的，而对于图像而言，可能那些 Curvelet 和 contourlet 是最优的，对于有些信号，也有可能需要将几种基结合起来才是最优的。稀疏分解是找到信号的最稀疏最有效的表达。

信号在某种表示方式下的稀疏性，是压缩感知应用的理论基础，经典的稀疏化的方法有离散余弦变换（DCT）、傅里叶变换（FFT）、离散小波变换（DWT）等。

最近几年，对稀疏表示研究的另一个热点是信号在冗余字典下的稀疏分解。这是一种全新的信号表示理论：用超完备的冗余函数库取代基函数，称之为冗余字典，字典中的元素被称为原子。目前信号在冗余字典下的稀疏表示的研究集中在两个方面：一是如何构造一个适合某一类信号的冗余字典，二是如何设计快速有效的稀疏分解算法。目前常用的稀疏分解算法大致可分为匹配追踪（Matching Pursuit）和基追踪（Basis Pursuit）两大类。

2、信号的观测矩阵

观测矩阵（也称测量矩阵） $M \times N$ ($M < N$) 是用来对 N 维的原信号进行观测得到 M 维的观测向量 Y ，然后可以利用最优化方法从观测值 Y 中大概率重构 X 。也就是说原信号 X 投影到这个观测矩阵（观测基）上得到新的信号表示 Y 。

观测矩阵的设计目的是如何采样得到 M 个观测值，并保证从中能重构出长度为 N 的信号 X 或者稀疏基 Ψ 下等价的稀疏系数向量。

为了保证能够从观测值准确重构信号，其需要满足一定的限制：观测基矩阵与稀疏基矩阵的乘积满足 RIP 性质（有限等距性质）。这个性质保证了观测矩阵不会把两个不同的 K 稀疏信号映射到同一个集合中（保证原空间到稀疏空间的一一映射关系），这就要求从观测矩阵中抽取的每 M 个列向量构成的矩阵是非奇异的。

在 CS 编码测量模型中并不是直接测量稀疏信号 X 本身，而是将信号投影到一组测量矩阵 Φ 上而得到测量值 y 。即，用一个与变换矩阵不相关的 $M \times N$ ($M < N$) 测量矩阵 Φ 对信号 x 进行线性投影，得到线性测量值 y ： $y = \Phi x$ ；

测量值 y 是一个 M 维向量，这样使测量对象从 N 维降为 M 维。测量矩阵的设计要求信号从 x 转换为 y 的过程中，所测量到的 K 个测量值不会破坏原始信号的信息，以保证信号可以精确重构。

由于信号 x 是可稀疏表示的： $x = \Psi s$ ，上式可以表示为下式：

$$y = \Phi x = \Phi \Psi s = \Theta s$$

其中 Φ 是一个 $M \times N$ 矩阵。上式中，方程的个数远小于未知数的个数，方程无确定解，无法重构信号。但是，由于信号是 K 稀疏，若上式中的 Φ 满足有限等距性质(Restricted Isometry Property, 简称 RIP)，则 K 个系数就能够从 M 个测量值准确重构（得到一个最优解）。RIP 性质的等价条件是测量矩阵 Φ 和稀疏基 Ψ 不相关。

如果稀疏基和观测基不相关，则很大程度上保证了 RIP 性。CandeS 和 Tao 等证明:独立同分布的高斯随机测量矩阵可以成为普适的压缩感知测量矩阵。则一般用随机高斯矩阵作为观测矩阵。目前常用的测量矩阵还有随机贝努利矩阵、部分正交矩阵、托普利兹和循环矩阵和稀疏随机矩阵等，这里不一一列举了。

3、信号的重构算法

当矩阵 Φ 满足 RIP 准则时。压缩感知理论能够通过对上式的逆问题先求解稀疏系数 s ，然后将稀疏度为 K 的信号 x 从 M 维的测量投影值 y 中正确地恢复出来。解码的最直接方法是通过 L_0 范数（ 0 -范数，也就是向量 y 中非零元素的个数）下求解的最优化问题：

$$\min_{\alpha} \|\alpha\|_{l_0} \quad s.t. \quad y = \Phi \Psi \alpha$$

从而得到稀疏系数 s 的估计 s' 。则原信号 $x' = \Psi s'$ 。由于上式的求解是个 NP 难问题（在多项式时间内难以求解，甚至无法验证解的可靠性）。 L_1 最小范数下在一定条件下和 L_0 最小范数具有等价性，可得到相同的解。那么上式转化为 L_1 最小范数下的最优化问题：

$$\min_{\alpha} \|\alpha\|_{l_1} \quad s.t. \quad y = \Phi \Psi \alpha$$

L_1 范数最小化是通过用 L_1 范数来近似 0 范数，取 1 而不取 $1/2, 2/3$ 或者其他值，是因为 1 范数最小化是凸优化问题，可以将求解过程转化成有一个线性规划问题。 L_1 最小范数下最优化问题又称为基追踪(BP)，其常用实现算法有：内点法和梯度投影法。内点法速度慢，但得到的结果十分准确；而梯度投影法速度快，但没有内点法得到的结果准确。

目前，压缩感知的重构算法主要分为两大类：

(1) 贪婪算法，它是通过选择合适的原子并经过一系列的逐步递增的方法实现信号矢量的逼近，此类算法主要包括匹配跟踪算法、正交匹配追踪算法、补空间匹配追踪算法等。

(2) 凸优化算法，它是把 0 范数放宽到 1 范数通过线性规划求解的，此类算法主要包括梯度投影法、基追踪法、最小角度回归法等。

凸优化算法比贪婪算法所求的解更加精确，但是需要更高的计算复杂度。

从数学上来说，CS 就是在一定的条件下求解欠定(不适定)方程，条件包括 x 要是稀疏的，测量矩阵要满足 RIP 条件，那么欠定(不适定)方程就会以很大的概率有唯一解。