

压缩感知综述

联系方式: qiuyunzou@qq.com

压缩感知理论由E.J.Candes、T.Tao、J.Romberg以及Donoho等科学家于2004年提出。压缩感知理论一经提出,便对信号处理领域产生了巨大的冲击。压缩感知理论作为一种新的采样原理,它通过以远小于Nyquist采样定律规定最小采样频率的前提下,用随机采样获取信号的离散样本,然后通过非线性重建算法完美的重建信号。压缩感知理论主要有两个原理性问题:其一,观测矩阵 Φ 需要满足什么条件,才能使得从观测样本中恢复出目标信号成为可能;其二,如何从观测向量中重构出目标信号。

1. 信号的稀疏分解

通常,大部分信号都具有稀疏特性。特别地,一个稠密信号亦可通过某个线性变换矩阵映射得到稀疏信号。信号的稀疏表示就是在给定的字典矩阵 Ψ 中用尽可能少的原子(或者称基,字典矩阵 Ψ 的列)的线性组合来表示信号,从而或者更为简洁的表达式,以去除信号的冗余信息。稀疏信号更加方便进一步加工处理,如压缩、编码等。信号的稀疏表示方向的研究热点主要集中在稀疏分解算法、过完备字典以及稀疏表示的应用等方面。

给定一个字典矩阵 Ψ ,它的每一列 $\{\psi_i\}_{i=1}^N$ 都是一个原子。信号 $\mathbf{x} \in \mathbb{R}^N$ 用字典矩阵 Ψ 的原子表示为

$$\mathbf{x} = \sum_{n=1}^N s_n \psi_n = \Psi \mathbf{s} \quad (1)$$

其中 Ψ 是一个 $N \times N$ 的正交矩阵。 $\mathbf{s} \in \mathbb{R}^N$ 为稀疏信号,其中 \mathbf{s} 中非零元素的个数 k (稀疏度)远小于信号长度 N ,即 $k \ll N$ 。实际应用中,也可将接近于0的值设置为0。信号的稀疏分解所用到的变换矩阵通常有DFT基、DCT基、小波基等。

■ DFT基

对于信号 $\mathbf{x} = [x_0, x_1, \dots, x_{N-1}]^T$,其傅里叶变换序列为 $\mathbf{X} = [X_0, X_1, \dots, X_{N-1}]^T$,两者之间的对应关系如下

$$X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n W_N^{kn} \quad (k = 0, \dots, N-1) \quad (2)$$

其中 $W_N = e^{-j\frac{2\pi}{N}}$ 称之为旋转因子。

通常,若信号为有限频率的信号线性叠加时,通过DFT基映射得到信号为稀疏信号。若信号中包含的频率成分较多时,信号在DFT基下通常不具有稀疏特性。通常将(2)用矩阵形式表示,即

$$\mathbf{X} = \mathbf{F}\mathbf{x} \quad (3)$$

其中

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{pmatrix} \quad (4)$$

很容易验证 \mathbf{F} 是正交矩阵。此外,通过对DFT基矩阵进行扩展得到过完备矩阵 $\mathbf{D} = [\mathbf{I}_N, \mathbf{F}_N]$ 。

■ DCT基

离散余弦变换 (discrete cosine transform, DCT) 是一种实数域变换。通过离散余弦变换,信号的重要信息可以集中到一小部分系数中,因此DCT稀疏基也是信号稀疏分解较为常用的一组稀疏基。对于长度 N 的离散时间列向量 $\mathbf{x} = [x_0, \dots, x_{N-1}]^T$,该信号对应的DCT变换序列为 $\mathbf{C} = [C_0, C_1, \dots, C_{N-1}]^T$,两者的对应关系为

$$C_k = a_k \sum_{n=0}^{N-1} x_n \cos\left(\frac{(2n+1)k\pi}{2N}\right) \quad k = 0, \dots, N-1 \quad (5)$$

其中

$$a_k = \begin{cases} \sqrt{\frac{1}{N}} & k = 0 \\ \sqrt{\frac{2}{N}} & \text{otherwise} \end{cases} \quad (6)$$

该公式(5)的矩阵形式表示为 $\mathbf{C} = \mathbf{F}\mathbf{x}$, 其中 \mathbf{F} 表示为DCT变换矩阵

$$\mathbf{F} = \begin{pmatrix} \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \cdots & \frac{1}{\sqrt{N}} \\ \sqrt{\frac{2}{N}} \cos\left(\frac{2\pi}{2N}\right) & \sqrt{\frac{2}{N}} \cos\left(\frac{3 \times 2\pi}{2N}\right) & \sqrt{\frac{2}{N}} \cos\left(\frac{5 \times 2\pi}{2N}\right) & \cdots & \sqrt{\frac{2}{N}} \cos\left(\frac{(2N-1) \times 2\pi}{2N}\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{\frac{2}{N}} \cos\left(\frac{(N-1)\pi}{2N}\right) & \sqrt{\frac{2}{N}} \cos\left(\frac{3(N-1)\pi}{2N}\right) & \sqrt{\frac{2}{N}} \cos\left(\frac{5(N-1)\pi}{2N}\right) & \cdots & \sqrt{\frac{2}{N}} \cos\left(\frac{(2N-1)(N-1)\pi}{2N}\right) \end{pmatrix}$$

可以验证DCT基亦为正交矩阵。

Remarks:

- 信号的稀疏表示,首先需要找到合适的基矩阵 Ψ 。将信号投影到该基矩阵张成的空间中 \mathcal{C} ,得到稀疏信号 \mathbf{s} ,该稀疏信号的几何意义为投影系数。
- 求解 \mathbf{s} ,可以理解为解线性方程组 $\mathbf{x} = \Psi\mathbf{s}$ 。压缩感知领域,求解稀疏信号的方法有贪婪迭代算法,如匹配追踪、正交匹配追踪等,详见稀疏信号重构小节。

2. 感知矩阵的条件

考虑无噪声情况,若 \mathbf{x} 为稠密信号,则观测向量表示为 $\mathbf{y} = \Phi\mathbf{x} = \Phi\Psi\mathbf{s}$,我们称 $\mathbf{A} = \Phi\Psi$ 为感知矩阵或传感矩阵。若 \mathbf{x} 为稀疏信号,则观测向量表示为 $\mathbf{y} = \Phi\mathbf{x}$,此时 $\mathbf{A} = \Phi$ 为传感矩阵,亦称观测矩阵。

从观测向量 \mathbf{y} 中重构出稀疏信号 \mathbf{x} ,要求观测矩阵 \mathbf{A} 满足一些特性。在无噪声情况下,零空间属性 (Null space property, NSP) 给出观测矩阵所需要满足的条件。然而实际系统中,观测向量不可避免会受到噪声的干扰。这种情况,欲从观测信号 \mathbf{y} 中恢复出稀疏信号 \mathbf{x} ,观测矩阵需要满足有限等距原理 (Restricted isometry property, RIP)。RIP的作用在于说明一个矩阵 \mathbf{A} 是否合适做观测矩阵,换言之,对于实际系统中,若感知矩阵 \mathbf{A} 满足有限等距原理,则可以大概率从观测向量中重构出稀疏信号。通常,随机矩阵满足RIP条件。

3. 信号重构

信号重构理论是压缩感知理论的核心,考虑模型

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \quad (7)$$

其中 $\mathbf{x} \in \mathbb{R}^N$ 表示稀疏信号, $\mathbf{H} \in \mathbb{R}^{M \times N}$ 表示观测矩阵, $\mathbf{y} \in \mathbb{R}^M$ 表示观测向量, $\mathbf{w} \in \mathbb{R}^M$ 为噪声干扰。与其他重构理论所不同,压缩感知中,观测向量的长度远小于目标信号的长度,即 $M \ll N$ 。由于目标信号 \mathbf{x} 为稀疏信号,设稀疏度为 k (目标信号中非零元素的个数),相应地,等价于求解 $M \times k$ 方程组。

这里主要介绍传统的压缩感知理论的重构算法:贪婪迭代算法,如匹配追踪算法、正交匹配追踪、子空间追踪等。

■ 匹配追踪算法

匹配追踪 (matching pursuit, MP) 算法最早用于时频分析,其目的是将已知信号拆分成许多原子信号 $\{\mathbf{h}_i\}_{i=1}^N$ (也称基信号,可非正交) 的线性组合,试图找到已知信号的近似表示。

$$\mathbf{y} = \sum_{i=1}^N a_i \mathbf{h}_i \quad (8)$$

其中 a_i 是权重, \mathbf{h}_i 表示原子或者基。

匹配追踪算法 (见表1),其收敛性可以通过残差更新的情况得到证明。更新之后的残差 ϵ^{t+1} 表示为 $\epsilon^{t+1} = \epsilon^t - \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle \mathbf{h}_{p[t]}$,设原子均经过列归一化,即 $\|\mathbf{h}_i\|_2 = 1$,则有

$$\|\epsilon^{t+1}\|_2 = \|\epsilon^t\|_2 - \left| \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle \right| \leq \|\epsilon^t\|_2 \quad (9)$$

因此,MP算法的残差是逐渐减小的。

Remarks:

- MP的思想是找到与残差 $\epsilon^t (\epsilon^0 = \mathbf{y})$ 内积最小的原子 $\mathbf{h}_{p[t]}$ 进行投影，投影系数为对应位置的 \hat{x} 估计值。随着每一次迭代的残差逐渐缩小达到收敛。然而，MP算法并不能保证当前残差与已经投影过得原子正交，换言之，当前残差仍有可能选择之前投影过的原子进行投影。
- 最终的表达式，观测向量表示为 $\mathbf{y} = \sum_{t=1}^T \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle \mathbf{h}_{p[t]} + \epsilon^{T+1}$ ，即观测向量被表示成观测矩阵 \mathbf{H} 的原子的线性组合。

表1 匹配追踪算法

Algorithm: Matching pursuit, MP
1. Initiation: residual error $\epsilon^0 = \mathbf{y}$ and estimator $\hat{\mathbf{x}} = 0$
2. Iteration ($t < T$)
a. Find the position of the best match autom of residual error.
$p^{[t]} = \arg \max_{i \in [N]} \langle \mathbf{h}_i, \epsilon^t \rangle $
b. Update residual error.
$\epsilon^{t+1} = \epsilon^t - \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle \mathbf{h}_{p[t]}$
c. Update estimator
$\hat{x}_{p[t]} = \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle$
c. set $t \leftarrow t + 1$ and proceed to step (a) until termination conditions are satisfied.
3. Output: $\mathbf{y} = \sum_{t=1}^T \langle \epsilon^t, \mathbf{h}_{p[t]} \rangle \mathbf{h}_{p[t]} + \epsilon^{T+1}$, and $\hat{\mathbf{x}}$.

■ 正交匹配追踪

前面提到，匹配追踪算法在迭代过程中不能保证残差与已投影过的原子的正交性。因此，在残差迭代过程中，仍有可能选择之前投影过的原子，从而产生不必要的迭代。正交匹配追踪算法的改进之处在于让残差与已选择过的原子正交，这样可以避免多次迭代，减少迭代次数。

正交匹配追踪通过使用最小二乘法（步骤c），使得残差与已匹配过的原子达到正交的目的。

$$\hat{\mathbf{x}}^{t+1} = \mathbf{H}_{S[t+1]}^\dagger \mathbf{y} \quad (10)$$

其中 $\mathbf{H}_{S[t+1]}^\dagger = \left(\mathbf{H}_{S[t+1]}^T \mathbf{H}_{S[t+1]} \right)^{-1} \mathbf{H}_{S[t+1]}$ 以及 $\mathbf{H}_{S[t+1]}$ 表示已选择过的原子构成的矩阵。为了解最小二乘的含义，我们通过如下图进行阐述。

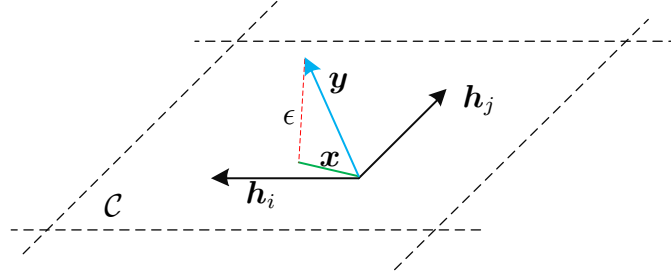


图1 最小二乘的几何解释

图中平面 C 是由 $H_{S[t+1]}$ 所张成的超平面。最小二乘的几何意义在于将 y 投影为该超平面 x , x 中的各元素所表示的是投影向量在各个原子上的系数。此时残差表示为

$$\epsilon^{t+1} = y - H_{S[t+1]} \hat{x}^{t+1} \quad (11)$$

直观上, 从图可以理解, 残差 ϵ^{t+1} 与 $H_{S[t+1]}$ 是正交的。

表2 正交匹配追踪

Algorithm: Orthogonal matching pursuit, OMP

1. Initiation: $\epsilon^0 = y$, $S^{[0]} = \emptyset$, $\hat{x} = 0$.

2. Iteration

a. Find the position of the best match atom of residual error

$$p^{[t]} = \arg \max_{i \in \mathcal{R} - S^{[t]}} |\langle h_i, \epsilon^t \rangle|$$

where $\mathcal{R} = \{1, \dots, N\}$, while $S^{[t]}$ denotes the set of column index of H that had been matched. $\mathcal{R} - S^{[t]}$ is the set of column index that don't be selected.

b. Update the index set

$$S^{[t+1]} = S^{[t]} \cup p^{[t]}$$

c. Estimator by LS

$$\hat{x}^{t+1} = H_{S[t+1]}^\dagger y$$

where $H_{S[t+1]}^\dagger = (H_{S[t+1]}^T H_{S[t+1]})^{-1} H_{S[t+1]}$ and $H_{S[t+1]}$ is from H indexed by $S^{[t+1]}$.

d. Update residual error

$$\epsilon^{t+1} = y - H_{S[t+1]} \hat{x}^{[t+1]}$$

3. Output: \hat{x}^{T+1} .

Remarks:

- 与MP让残差投影到最匹配原子方式所不同, OMP的做法是将残差投影到已选择原子所构成的超平面上, 从而达到残差与已匹配的原子正交的目的。
- 由于OMP减少了残差在已匹配原子上投影, 其收敛速度相对于MP算法要快, 其代价是增加每一次迭代次数的计算复杂度。OMP的计算复杂度主要集中在矩阵求逆上, 即步骤c中 (矩阵求逆的计算量大概是矩阵维度的3次方)。
- OMP算法的收敛性可以对比两次残差的范数进行证明。

■ 子空间追踪

子空间追踪 (Subspace pursuit, SP) 算法与OMP很类似。都是将 \mathbf{y} 投影到索引集合 $\mathcal{K}^{[t]}$ 所索引的原子构成的超平面上, 所不同的在于寻找 $\mathcal{K}^{[t]}$ 的方式不同结果。此外, 子空间追踪还需要依赖信号稀疏度这个先验条件。SP算法详见表3。

表3 子空间追踪

Algorithm: Subspace pursuit, SP
1. Initiation: $\epsilon^0 = \mathbf{y}$, $\mathcal{S}^{[0]} = \emptyset$.
2. Iteration
a. Find the index set $\mathcal{C}^{[t]}$ of the best match k atoms of residual error in descend order and then update $\mathcal{S}^{[t]} = \mathcal{S}^{[t-1]} \cup \mathcal{C}^{[t]}$.
b. Calculate $\mathbf{x}_{\text{tem}} = (\mathbf{H}_{\mathcal{S}^{[t]}}^T \mathbf{H}_{\mathcal{S}^{[t]}})^{-1} \mathbf{H}_{\mathcal{S}^{[t]}} \mathbf{y}$.
c. Find the index set $\mathcal{K}^{[t]}$ represented the largest absolute k elements in \mathbf{x}_{tem} and set $\mathcal{S}^{[t]} = \mathcal{K}^{[t]}$.
d. Estimation and update residual error: $\tilde{\mathbf{x}}^t = (\mathbf{H}_{\mathcal{K}^{[t]}}^T \mathbf{H}_{\mathcal{K}^{[t]}})^{-1} \mathbf{H}_{\mathcal{K}^{[t]}} \mathbf{y}$ $\epsilon^{t+1} = \mathbf{y} - \mathbf{H}_{\mathcal{S}^{[t]}} \hat{\mathbf{x}}^{[t]}$
3. Output: $\hat{\mathbf{x}}_{\mathcal{K}^{[t]}} = \tilde{\mathbf{x}}^t$, $\hat{\mathbf{x}}_{\mathcal{R}-\mathcal{K}^{[t]}} = \mathbf{0}$ with $\mathcal{R} = \{1, \dots, N\}$.

Remarks:

- 子空间追踪, 顾名思义, 先找到子空间 $\mathcal{S}^{[t]}$ (步骤a), 然后从子空间中挑选出最大的 k 个值的集合 $\mathcal{K}^{[t]}$ 。相应地, 观测向量 \mathbf{y} 在 $\mathbf{H}_{\mathcal{K}^{[t]}}$ 上的投影向量, 对应于估计器的 $\hat{\mathbf{x}}_{\mathcal{K}^{[t]}}$ 。