Sparse representation of images

Jayanta Mukhopadhyay
Dept. of Computer Science and Engg.

Image Transform

$$f(x,y) = \sum_j \sum_i \lambda_{ij} b_{ij}(x,y)$$

- Image in continuous form: $f(x,y)$: A 2-D function, where (x,y) in R^2.
- Let B be a set of basis functions:

Properties of basis functions can be extended in the analysis.

$$B=\{b_i(x,y) \mid i=..,-1,0,1,2,3,\ldots.\},\ b_i(x,y) \text{ in } R \text{ or } C.$$

- Let $f(x,y)$ be expanded using B as follows:

$$f(x,y) = \sum_i \lambda_i b_i(x,y)$$

Coefficients of transform

The **transform** of f w.r.t. B is given by $\{\lambda_i | i = \cdots.-1,0,1,2,3,\ldots\}$.

Indexing may be multidimensional say, λ_{ij}.

Orthogonal Expansion and 1-D Transforms

$$f(x) = \sum_i \lambda_i b_i(x)$$

- Inner product: $$< f, g >= \int f(x) g^*(x) dx$$

- Orthogonal expansion: If B satisfies :

$$\begin{aligned} < b_i.b_j > &= 0, \quad for\ i \neq j, \\ &= c_i, \quad Otherwise\ (for\ i = j),\ where\ c_i > 0. \end{aligned}$$

comma

- Transform coefficients in O.E.: $\lambda_i = \frac{1}{c_i} < f.b_i >$
- If c_i=1, it becomes orthonormal expansion.

Forward transform $\lambda_i =< f.b_i >$

- Inverse transform: $$f(x) = \int_{i=-\infty}^{\infty} \lambda_i . b_i(x) di$$

Fourier transform

Complete base

$$B = \{e^{-j\omega x} | -\infty < \omega < \infty\}$$

Orthogonality:

$$\int_{-\infty}^{\infty} e^{j\omega x} dx = \begin{cases} 2\pi\delta(\omega), & \text{for } \omega = 0, \\ 0, & \text{Otherwise.} \end{cases}$$

Fourier Transform:

$$\mathbb{F}(f(x)) = \hat{f}(j\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx$$

Inverse Transform:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(j\omega).e^{j\omega x} d\omega$$

Full reconstruction

$$e^{-j\omega x} = \cos(\omega x) - j\sin(\omega x)$$

$$\hat{f}(j\omega) = \int_{-\infty}^{\infty} f(x)(\cos(\omega x) - j\sin(\omega x))\, dx$$

$$C = \{\cos(\omega x) | -\infty < \omega < \infty\} \qquad S = \{\sin(\omega x) | -\infty < \omega < \infty\}$$

Orthogonal

But not complete!

Discrete representation

- Discrete representation of a function:

$$f(n) = \{f(nX_0) | n \in \mathbb{Z}\}$$

Set of integers

Sampling interval

- Can be considered as a vector in an infinite dimensional vector space.
- In our context, it is of a finite dimensional space, e.g. $\{f(n), n=0,1,..N\text{-}1\}$, or
- $\boldsymbol{f}=[f(0)\ f(1)\ \ldots\ f(N\text{-}1)]^T$.

Discrete Transform

- For n-dimensional vector X any linear transform, e.g. $Y_{m\text{x}1}=B_{m\text{x}n}X_{n\text{x}1}$
- Has inverse transform if B is square of size (nxn) and invertible.
- Rows of B are called basis vectors.
- $Y(i)=<\boldsymbol{b}_i^{*T}. X>$

dot product or inner product.

$$B = \begin{bmatrix} \boldsymbol{b}_0^{*T} \\ \boldsymbol{b}_1^{*T} \\ \vdots \\ \boldsymbol{b}_n^{*T} \end{bmatrix}$$

- Orthogonality condition:

$$\begin{aligned} <\boldsymbol{b}_i^{*T}.\boldsymbol{b}_j> &= 0 \; if \; i \neq j \\ &= c_i, \qquad otherwise \end{aligned}$$

Discrete Fourier Transform (DFT)

$$b_k(n) = \frac{1}{\sqrt{N}} e^{j2\pi \frac{k}{N} n}, \text{ for } 0 \le n \le N-1, \text{ and } 0 \le k \le N-1.$$

$$\hat{f}(k) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi \frac{k}{N} n} \text{ for } 0 \le k \le N-1.$$

$$\hat{f}(N+k) = \hat{f}(k)$$

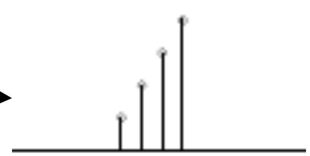
k/N: Normalized frequency

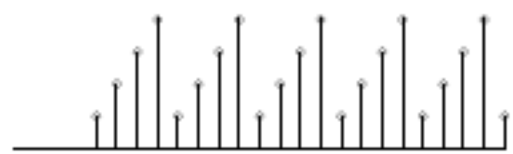
$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}(k) e^{j2\pi \frac{k}{N} n} \text{ for } 0 \le n \le N-1.$$

Hermitian Transpose

$$\mathbf{X} = \mathbf{F}\mathbf{x} \qquad \mathbf{F} = \left[e^{-j2\pi \frac{k}{N} n} \right]_{0 \le (k,n) \le N-1}. \qquad \mathbf{F}^{-1} = \frac{1}{N} \mathbf{F}^H$$

A single period





$f(n+N)=f(n)$

DFT: Fourier series of a periodic function

Fundamental frequency: $1/(NX_0)$

Generalized Discrete Fourier Transform (GDFT)

$$\mathbf{F}_{\alpha,\beta} = \left[e^{-j2\pi \frac{k+\alpha}{N}(n+\beta)} \right]_{0 \le (k,n) \le N-1}.$$

$$\begin{aligned}
\mathbf{F}_{0,0}^{-1} &= \tfrac{1}{N}\mathbf{F}_{0,0}^{H} &= \tfrac{1}{N}\mathbf{F}_{0,0}^{*}, \\
\mathbf{F}_{\frac{1}{2},0}^{-1} &= \tfrac{1}{N}\mathbf{F}_{\frac{1}{2},0}^{H} &= \tfrac{1}{N}\mathbf{F}_{0,\frac{1}{2}}^{*}, \\
\mathbf{F}_{0,\frac{1}{2}}^{-1} &= \tfrac{1}{N}\mathbf{F}_{0,\frac{1}{2}}^{H} &= \tfrac{1}{N}\mathbf{F}_{\frac{1}{2},0}^{*}, \text{ and} \\
\mathbf{F}_{\frac{1}{2},\frac{1}{2}}^{-1} &= \tfrac{1}{N}\mathbf{F}_{\frac{1}{2},\frac{1}{2}}^{H} &= \tfrac{1}{N}\mathbf{F}_{\frac{1}{2},\frac{1}{2}}^{*}.
\end{aligned}$$

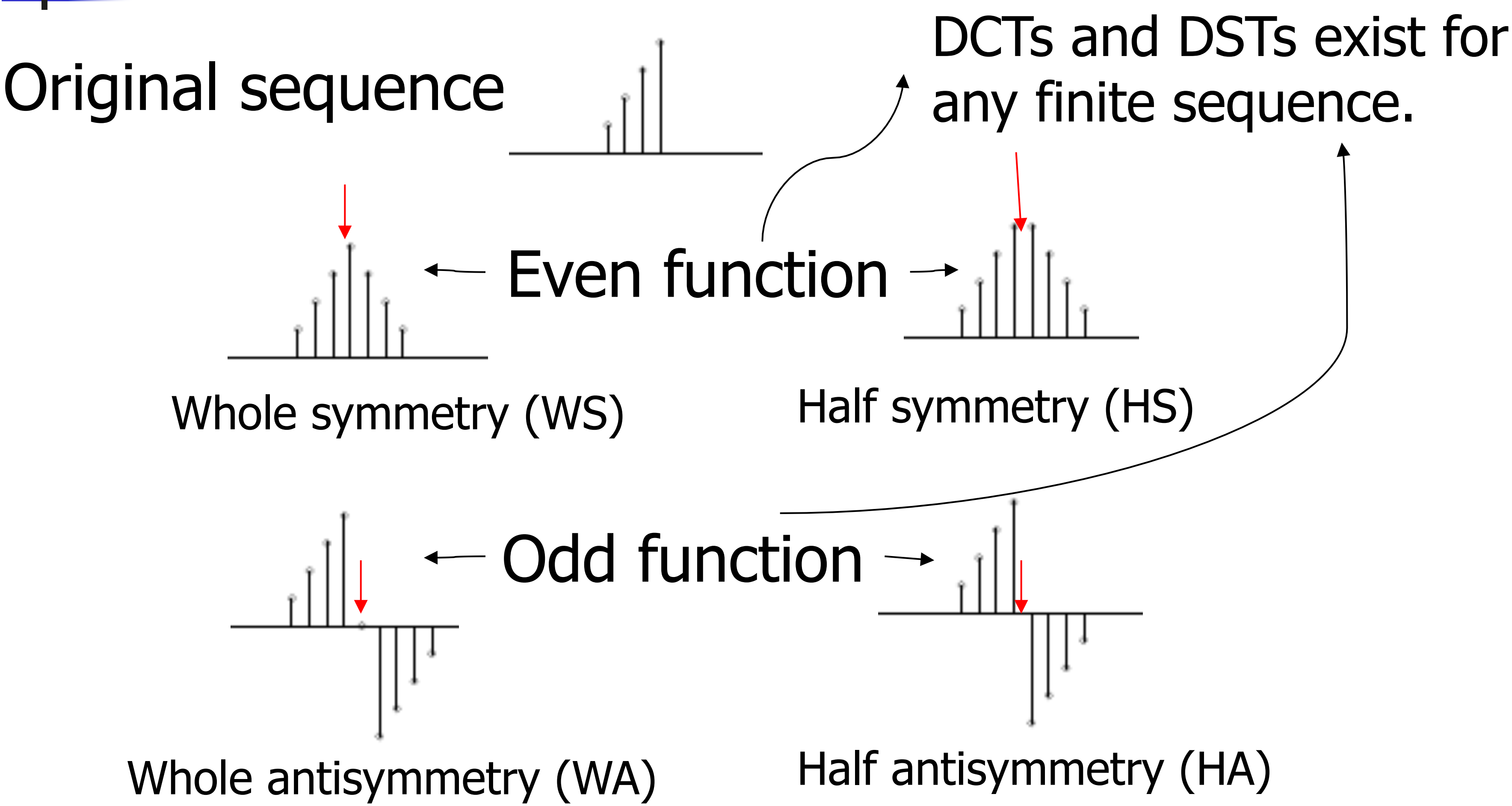
$$b_k^{(\alpha,\beta)}(n) = \frac{1}{\sqrt{N}} e^{j2\pi \frac{k+\alpha}{N}(n+\beta)}, \text{ for } 0 \le n \le N-1, \text{ and } 0 \le k \le N-1$$

$$\hat{f}_{\alpha,\beta}(k) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi \frac{k+\alpha}{N}(n+\beta)}, \text{ for } 0 \le k \le N-1$$

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_{\alpha,\beta}(k) e^{j2\pi \frac{k+\alpha}{N}(n+\beta)}, \text{ for } 0 \le n \le N-1$$

α	β	Transform name	Notation
0	0	Discrete Fourier Transform (DFT)	$\hat{f}(k)$
0	$\frac{1}{2}$	Odd Time Discrete Fourier Transform ($OTDFT$)	$\hat{f}_{0,\frac{1}{2}}(k)$
$\frac{1}{2}$	0	Odd Frequency Discrete Fourier Transform ($OFDFT$)	$\hat{f}_{\frac{1}{2},0}(k)$
$\frac{1}{2}$	$\frac{1}{2}$	Odd Frequency Odd Time Discrete Fourier Transform (O^2DFT)	$\hat{f}_{\frac{1}{2},\frac{1}{2}}(k)$

Symmetric / Antisymmetric extension of a finite sequence

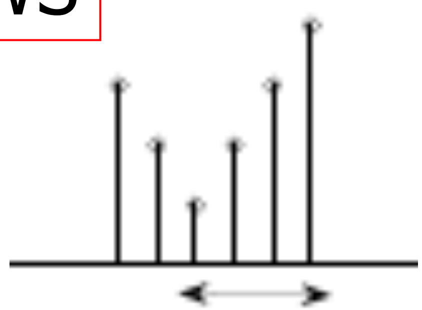


Discrete Cosine / Sine Transforms

$$\alpha(p) = \begin{cases} \sqrt{\frac{1}{2}}, & \text{for } p = 0 \text{ or } N, \\ 1, & \text{otherwise.} \end{cases}$$

- Types of symmetric / antisymmetric extensions at the two ends of a sequence and a type of GDFT→ DCTs / DSTs

WSWS



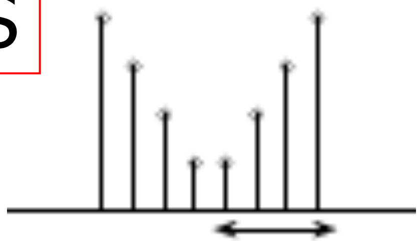
$\xrightarrow{F_{0,0}}$ Type-I Even DCT

$$C_{1e}(x(n)) = X_{Ie}(k) = \sqrt{\frac{2}{N}}\alpha^2(k)\sum_{n=0}^{N} x(n)\cos\left(\frac{2\pi nk}{2N}\right),\ 0 \le k \le N;$$

Discrete Cosine / Sine Transforms

$$\alpha(p) = \begin{cases} \sqrt{\frac{1}{2}}, & \text{for } p = 0 \text{ or } N, \\ 1, & \text{otherwise.} \end{cases}$$

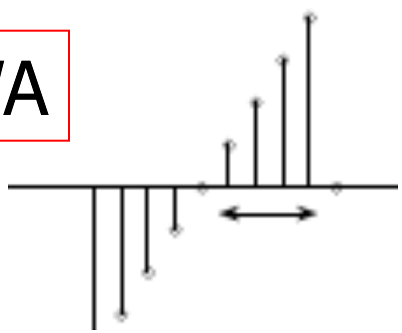
HSHS



$\xrightarrow{F_{0,1/2}}$ Type-2 Even DCT

$$C_{2e}(x(n)) = X_{IIe}(k) = \sqrt{\frac{2}{N}}\alpha(k)\sum_{n=0}^{N-1} x(n)\cos\left(\frac{2\pi k(n+\frac{1}{2})}{2N}\right),\ 0 \le k \le N-1$$

WAWA



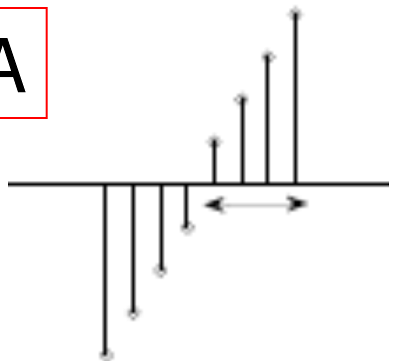
$\xrightarrow{jF_{0,0}}$ Type-1 Even DST

$$S_{1e}(x(n)) = X_{sIe}(k) = \sqrt{\frac{2}{N}}\sum_{n=1}^{N-1} x(n)\sin\left(\frac{2\pi kn}{2N}\right),\ 1 \le k \le N-1$$

Discrete Cosine / Sine Transforms

$$\alpha(p) = \begin{cases} \sqrt{\frac{1}{2}}, & \text{for } p = 0 \text{ or } N, \\ 1, & \text{otherwise.} \end{cases}$$

HAHA



$$\xrightarrow{jF_{0,1/2}}$$

Type-2 Even DST

$$S_{2e}(x(n)) = X_{sIIe}(k) = \sqrt{\frac{2}{N}}\alpha(k)\sum_{n=0}^{N-1} x(n)\sin\left(\frac{2\pi k(n+\frac{1}{2})}{2N}\right), \; 1 \le k \le N-1$$

There exist 16 different types of DCTs and DSTs.

Type-II Even DCT is used in signal, image, and video compression.

$$f(x,y) = \sum_j \sum_i \lambda_{ij} b_{ij}(x,y)$$

2-D Transforms

- Easily extendable if basis functions are separable, i.e. $B=\{\ b_{ij}(x,y)=g_i(x).g_j(y)\}$.

They could be from two different sets, say $b(x,y)=g(x).h(y)$.

1-D basis function

- B: Orthogonal if $G=\{g_i(x),\ i=1,2,..\}$ is orthogonal.
- B: Orthogonal and complete if G is so.
- Reuse of 1-D transform computation.

$$\lambda_{ij} = \sum_j g_j^*(y) \left(\sum_i f(x,y) g_i^*(x) \right)$$

2D Discrete Transform

$$Y_{mxn} = B_{mxm} X_{mxn} B_{nxn}^{\mathrm{T}}$$

- Use of separability:
  - Transform columns.
  - Transform rows.
- Input: X_{mxn} 1-D Transform Matrix: *B*
- Transform columns: $[Y_1]_{mxn} = B_{mxm} X_{mxn}$
- Transform rows: $Y_{mxn} = [B_{nxn} Y_1^{\mathrm{T}}]^{\mathrm{T}}$

$$= Y_1 B_{nxn}^{\mathrm{T}}$$

$$= B_{mxm} X_{mxn} B_{nxn}^{\mathrm{T}}$$

Sparse representation

- Some signals cannot be represented efficiently in an orthonormal basis.
  - intermixture of impulses and sinusoids
    - Inefficient to represent by only impulses or only sinusoids
- Use of redundant set of basis functions (Dictionary)
  - A Gabor dictionary
    - complex exponentials smoothly windowed to short time intervals used for joint time–frequency analysis.
- Best linear combination of elements of redundant dictionary.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

Nonlinear approximation

- Projection onto the best linear subspace spanned by *m* elements of a fixed orthonormal basis.
- Redundant set of basis functions (Dictionary)
  - A Gabor dictionary: complex exponentials smoothly windowed to short time intervals used for joint time–frequency analysis.
- The problem of approximating a signal with the best linear combination of elements from a redundant dictionary is called sparse approximation or highly nonlinear approximation.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

Sparse Approximation: Problem statement

- The problem of approximating a signal with the best linear combination of elements from a redundant dictionary.
  - Optimal / Near optimal representation
  - Fast computation
  - Optimal dictionary (joint optimization problem)
- Two major approaches
  - Basis pursuit (BP)
  - Orthogonal Matching pursuit (OMP)

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

OMP

- An iterative greedy algorithm
  - selects at each step the dictionary element best correlated with the residual part of the signal.
  - produces a new approximant by projecting the signal onto the dictionary elements that have already been selected.
  - extends the trivial greedy algorithm that succeeds for an orthonormal system.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

BP

- A more sophisticated approach that replaces the original sparse approximation problem by a linear programming problem.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

EXACT-SPARSE problem

- Dictionary of *N* elementary *n-D* signals called atoms.
- To identify the representation of the input signal that uses the least number of atoms, i.e., the sparsest one.
- Given an input *n-D* signal *X*, form a matrix A_{opt} whose columns are the atoms that make up the optimal representation of the signal.
- The sparsest representation: $Y=A_{opt}^{+}X$.
  - Pseudo inverse: $A_{opt}^{+}=(A_{opt}^{*}A_{opt})^{-1}A_{opt}$
- Unique sparsest representation if $\max_{\psi\in B}\left\|A_{opt}^{+}\psi\right\|_1 \leq 1$
  - *B* is the complementary of A_{opt} in the dictionary.
- Both BP and OMP provides the same optimal solution.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

EXACT-SPARSE: Condition

- Unique sparsest representation if $\max_{\psi \in B} \left\| A_{opt}^{+} \psi \right\|_1 \leq 1$
  - B is the complementary of A_{opt} in the dictionary.
- Coherence parameter μ : the maximum absolute inner product between two distinct atoms.
- $\mu_1(m)$= the maximum absolute sum of inner products between a fixed atom and other atoms.
- m-term approximation exists if
  - $(m<(\mu^{-1} + 1)/2) \Rightarrow (\mu_1(m)+\mu_1(m-1)<1)$
- SPARSE problem: Minimize the approximation error using L_2 norm using m terms.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

SPARSE approximation: A variant

- SPARSE problem: Minimize the approximation error using L_2 norm using *m* terms.

Dictionary

$$\mathcal{D} = \{\varphi_\omega : \omega \in \Omega\}.$$

Optimization task

Linear combination

$$\min_{|\Lambda|=m} \min_{\{b_\lambda\}} \left\| s - \sum_{\lambda \in \Lambda} b_\lambda \varphi_\lambda \right\|_2$$

signal

Fixed no. of atoms.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

Matching pursuit

$$\mathcal{D} = \{\varphi_\omega : \omega \in \Omega\}.$$

$$\min_{|\Lambda|=m} \min_{\{b_\lambda\}} \left\| s - \sum_{\lambda\in\Lambda} b_\lambda \varphi_\lambda \right\|_2$$

- Minimize the approximation error using L_2 norm using *m* terms.

Residue (r_k) and Approximation (a_k)

Initialization $r_0 = s \quad a_0 = 0$

At *k* th step:

$$\lambda_k \in \arg\max_{\omega\in\Omega} \langle r_{k-1}, \varphi_\omega \rangle .$$

$$a_k = a_{k-1} + \langle r_{k-1}, \varphi_{\lambda_k} \rangle \varphi_{\lambda_k}$$

$$r_k = r_{k-1} - \langle r_{k-1}, \varphi_{\lambda_k} \rangle \varphi_{\lambda_k}$$

$$\Longrightarrow \quad r_k = s - a_k$$

MP may select the same atom multiple times.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

$$\mathcal{D} = \{\varphi_\omega : \omega \in \Omega\}.$$

OMP

$$\min_{|\Lambda|=m} \min_{\{b_\lambda\}} \left\| s - \sum_{\lambda\in\Lambda} b_\lambda \varphi_\lambda \right\|_2$$

- Minimize the approximation error using *L_2* norm using *m* terms.

Initialization

$$r_0 = s \quad a_0 = 0$$

At *k* th step:

$$\lambda_k \in \arg\max_{\omega\in\Omega} \langle r_{k-1}, \varphi_\omega \rangle .$$

$$a_k \stackrel{\text{def}}{=} \arg\min_{a} \|s - a\|_2 \quad \text{subject to } a \in \operatorname{span}\{\varphi_\lambda : \lambda \in \Lambda_k\}.$$

$$r_k = s - a_k$$

This minimization can be performed incrementally with standard least-squares techniques.

OMP may select an atom only once, as the residual is always orthogonal to selected set.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

$$\mathcal{D} = \{\varphi_\omega : \omega \in \Omega\}.$$

BP

Convex relaxation of EXACT-SPARSE problem.

- Minimize the approximation error using L_1 norm of m coefficients.
  - A convex function, hence can be minimized in polynomial time

$$\min_{\{b_\omega\}} \quad \sum_{\omega\in\Omega} |b_\omega| \quad \textbf{subject to} \quad \sum_{\omega\in\Omega} b_\omega \varphi_\omega = \boldsymbol{s}$$

Use linear programming.

Joel A. Tropp, , Greed is Good: Algorithmic Results for Sparse Approximation, IEEE TRANSACTIONS ON INFORMATION THEORY, 50 (10), 2004, 2231-2242.

Problem statement: sparse representation

- n-D signal : y in R^n
- K number of atoms.
- $n < K$
- Dictionary Matrix: D in R^{nxK} : $[d_j]^K_{j=1}$ Columns
- To obtain a sparse X in R^K such that
  - $Y=DX$, or $Y \sim DX$

$$\boxed{\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x}}$$

$$\boxed{\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon}$$

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

K-SVD: Forming dictionary for sparse representation

- Given a set of training signals $\{y_i\}_{i=1}^{N}$, to obtain the dictionary of K elements that leads to the best possible representations for each member in this set with strict sparsity constraints.

$$\boxed{\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x}}$$

$$\boxed{\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon}$$

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\mathbf{x}$$

K-SVD

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon$$

- Generalizes K-means clustering problem.
  1. Choose a dictionary of K atoms.
  2. Obtain sparse representation.
  3. Update dictionary atoms.
  4. Repeat steps 2 and 3 till convergence.
- K-means clustering: Extreme sparse representation of a signal by a single atom only.
- K-SVD: A sparse linear combination of K atoms.

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

K-means clustering

- Given a set of atoms $D=\{d_i\}_1^K$
  - Assign the training examples $\{y_i\}_{i=1}^N$ to their nearest neighbor in *D.*
    - *Usually L_2 norm used.*
  - Given the assignment update *D* to better fit the examples.
    - *Update mean of each partition of assignment.*
- *Start with any initial set of distinct atoms.*

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

K-means clustering: A code book with extreme sparse representation

- The code book: $D=\{d_i\}_1^K=[d_1\ d_2\ \ldots\ d_K\ ]_{n\mathrm{x}K}$
- The training examples: $[\,Y\,]_{n\mathrm{x}N}=\{y_i\}_{i=1}^N$
- Extreme sparse vector: $e_j=[\ 0\ 0\ \ldots 0\ 1\ 0\ \ldots\ 0]^T$
  - Only j th term is 1 of K-dim. vector.
- Sparse representation: $X=[x_1\ x_2\ \ldots\ x_N]_{K\mathrm{x}N}$
  - where x_i is one of e_j's .

Frobenius norm

- Optimization problem: Minimize $\|Y-DX\|^2_F$
  - $x_i=e_r$ if $\|y_i-d_r\|2$ is minimum among all atoms.
- Update atoms: $d_j=$ Mean $(\{y_i \mid x_i=e_j\})$, for all j.

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

KSVD: Generalization of K-means clustering

- The code book: $D=\{d_i\}_1^K=[d_1\ d_2\ \ldots\ d_K\ ]_{\mathrm{nxK}}$
- The training examples: $[\,Y\,]_{\mathrm{nxN}}=\{y_i\}_{i=1}^N$
- Sparse representation: $X=[x_1\ x_2\ \ldots\ x_N]_{\mathrm{KxN}}$
  - where x_i provides linear combination of maximum T_0 *nonzero* terms.
- Optimization problem:
  - Minimize $\|Y - DX\|^2_F$ subject to $\|x_i\|_0 \leq T_0$, for all *i*.

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

Minimize $||Y - DX||^2_F$ subject to $||x_i||_0 \leq T_0$, for all i.

Rewriting optimization function

- x^j_T: j th row of X.

Consider effect of minimizing w.r.t. k th row of X associated with code vector d_k keeping other terms fixed.

$$\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 = \left\| \mathbf{Y} - \sum_{j=1}^{K} \mathbf{d}_j \mathbf{x}_T^j \right\|_F^2$$

$$= \left\| \left(\mathbf{Y} - \sum_{j \neq k} \mathbf{d}_j \mathbf{x}_T^j \right) - \mathbf{d}_k \mathbf{x}_T^k \right\|_F^2$$

$$= \|\mathbf{E}_k - \mathbf{d}_k \mathbf{x}_T^k\|_F^2.$$

But the column vector may not be sparse.

Perform SVD: $E_k = UDV^T$ and take columns of U and V for max singular values.

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

Enforcing sparsity

$$\|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 = \left\|\mathbf{Y}-\sum_{j=1}^{K}\mathbf{d}_j\mathbf{x}_T^j\right\|_F^2 = \left\|\left(\mathbf{Y}-\sum_{j\neq k}\mathbf{d}_j\mathbf{x}_T^j\right)-\mathbf{d}_k\mathbf{x}_T^k\right\|_F^2 = \|\mathbf{E}_k-\mathbf{d}_k\mathbf{x}_T^k\|_F^2.$$

- x^j_T : *j* th row of *X*.
- Choose only samples from Y which have a nonzero component along d_k.
- Form reduced E_k (denoted E_{kR}) and x^k_T by x^k_R.
- Perform SVD of E_{kR} to get d_k and x^k_R.
- Update d_k and x^k_T.
- Repeat for all d_j's and obtain updated D and X.
- Repeat till convergence

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

The Algorithm

Task: Find the best dictionary to represent the data samples $\{\mathbf{y}_i\}_{i=1}^N$ as sparse compositions, by solving

$$\min_{\mathbf{D},\mathbf{X}} \{\|\mathbf{Y} - \mathbf{DX}\|_F^2\} \quad \text{subject to} \quad \forall i, \ \|\mathbf{x}_i\|_0 \le T_0.$$

Initialization : Set the dictionary matrix $\mathbf{D}^{(0)} \in \mathbb{R}^{n\times K}$ with ℓ^2 normalized columns. Set $J = 1$.

Repeat until convergence (stopping rule):

- *Sparse Coding Stage*: Use any pursuit algorithm to compute the representation vectors $\mathbf{x}_i$ for each example $\mathbf{y}_i$, by approximating the solution of

$$i = 1, 2, \ldots, N, \quad \min_{\mathbf{x}_i} \{\|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_2^2\} \quad \text{subject to} \quad \|\mathbf{x}_i\|_0 \le T_0.$$

- *Codebook Update Stage*: For each column $k = 1, 2, \ldots, K$ in $\mathbf{D}^{(J-1)}$, update it by
  - Define the group of examples that use this atom, $\omega_k = \{i \mid 1 \le i \le N, \ \mathbf{x}_T^k(i) \ne 0\}$.
  - Compute the overall representation error matrix, $\mathbf{E}_k$, by

$$\mathbf{E}_k = \mathbf{Y} - \sum_{j\ne k} \mathbf{d}_j \mathbf{x}_T^j.$$

  - Restrict $\mathbf{E}_k$ by choosing only the columns corresponding to ω_k, and obtain $\mathbf{E}_k^R$.
  - Apply SVD decomposition $\mathbf{E}_k^R = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^T$. Choose the updated dictionary column $\tilde{\mathbf{d}}_k$ to be the first column of $\mathbf{U}$. Update the coefficient vector $\mathbf{x}_R^k$ to be the first column of $\mathbf{V}$ multiplied by $\boldsymbol{\Delta}(1,1)$.
- Set $J = J + 1$.

Michal Aharon, Michael Elad, and Alfred Bruckstein, K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 11, NOVEMBER 2006, 4311-4322

Thank you!