Segmentation of Subspace Arrangements I – Introduction

Allen Y. Yang

Berkeley CS 294-6, Lecture 23

Nov. 20, 2006

1. What is a subspace arrangement
   - Definition
   - Representation

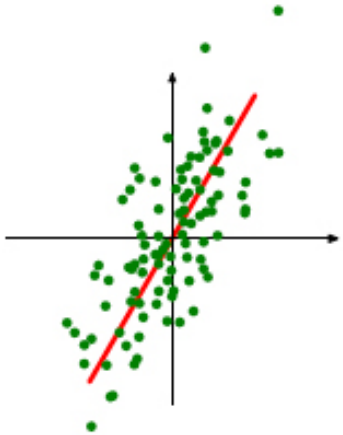
2. Converting constraints to subspace arrangements
   - Vanishing points
   - Affine projections
   - Perspective projections
   - Approximate nonlinear structures

3. Statistical solutions
   - K-Means and K-Subspaces
   - EM
   - RANSAC

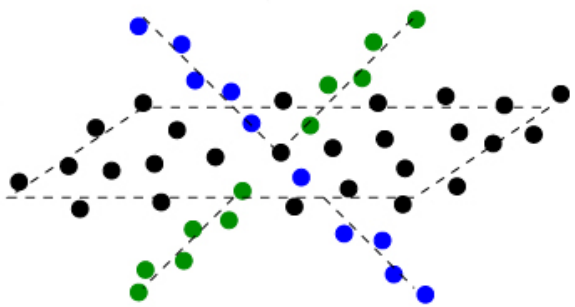
Linear Subspaces and Arrangements

1. Linear subspaces
   - Subspace structures are widely used for dimension reduction in pattern recognition.
   - Linear subspaces pass through the origin, contrary to affine subspaces.
   - The classic solution for estimating a subspace *for dimension reduction*: PCA
     Solution: Singular Value Decomposition (SVD)
     $U\Sigma V^T = \text{svds}([\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_N]_{D\times N}, d)$, where $U \in \mathbb{R}^{D\times d}$, $\Sigma \in \mathbb{R}^{d\times d}$, and $V \in \mathbb{R}^{N\times d}$.

     - $U(:, i)$ is the *ith principal axis* (basis vector).
     - $c_{ij} = U(:, i)^T \mathbf{z}_j$ is the *ith principal component of* $\mathbf{z}_j$.
     - The projected samples are
       $$[\mathbf{z}'_1, \mathbf{z}'_2, \cdots, \mathbf{z}'_N] = \Sigma V^T \in \mathbb{R}^{d\times N}.$$

2. A subspace arrangement is a collection of linear subspaces in $\mathbb{R}^D$, parametized by the number of subspaces K and their dimensions $d_1, d_2, \cdots, d_K$.



Description of Subspace Arrangements

1. Given one subspace $V \in \mathbb{R}^D$, $\dim(V) = d$, it is defined by d basis vectors:

$$V = \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_d)$$

Now $\{\mathbf{v}_1, \cdots, \mathbf{v}_d\}$ can be expanded to a complete basis of $\mathbb{R}^D$:

$$\{\mathbf{v}_1, \cdots, \mathbf{v}_d, \mathbf{u}_1, \cdots, \mathbf{u}_r\} \text{ where } d + r = D.$$

For any $\mathbf{z} \in V$, $\mathbf{z} \perp V^\perp \doteq \text{Span}(\mathbf{u}_1, \cdots, \mathbf{u}_r)$:

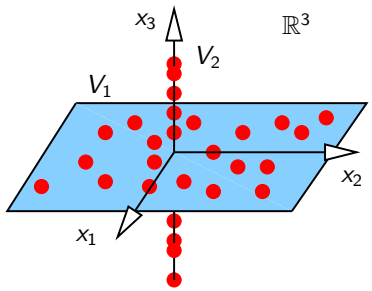
$$\forall \mathbf{z} \in V, (\mathbf{u}_1^T \mathbf{z} = 0) \wedge (\mathbf{u}_2^T \mathbf{z} = 0) \wedge \cdots \wedge (\mathbf{u}_r^T \mathbf{z} = 0).$$

2. With multiple subspaces $V_1, V_2, \cdots, V_K$. The arrangement is:

$$\mathcal{A} = V_1 \cup V_2 \cup \cdots \cup V_K.$$

Then for any $\mathbf{z} \in \mathcal{A}$, $(\mathbf{z} \in V_1) \vee (\mathbf{z} \in V_2) \vee \cdots \vee (\mathbf{z} \in V_K)$:

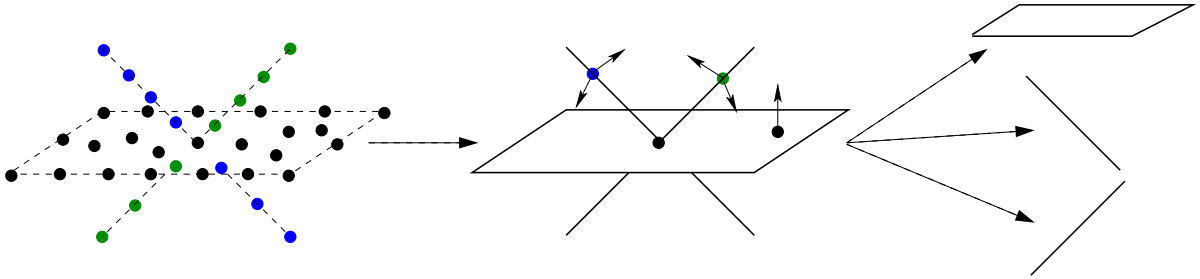
$$\forall \mathbf{z} \in \mathcal{A}, ({V_1^\perp}^T \mathbf{z} = 0) \vee ({V_2^\perp}^T \mathbf{z} = 0) \vee \cdots \vee ({V_r^\perp}^T \mathbf{z} = 0).$$



Punch line

Subspace arrangements are generalization of single subspaces. They are also more compact representation for dimention reduction.

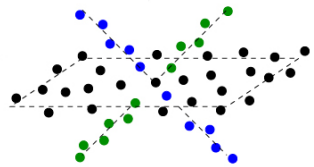
Segmentation of subspace arrangements (subspace-segmentation problem):
Given number of subspaces K and their dimensions $d_1, \cdots, d_K$ known,



A closer look: There are couples of problems

1. Polynomial interpretation:
   - $p_j(\mathbf{z}) \doteq \mathbf{u}_j^T \mathbf{z}_i$ is a 1st degree homogeneous polynomial with coefficients $\mathbf{u}_j$.
   - For a single V, $p_1(\mathbf{z}) = 0, \cdots, p_r(\mathbf{z}) = 0$.
2. Noise issue:

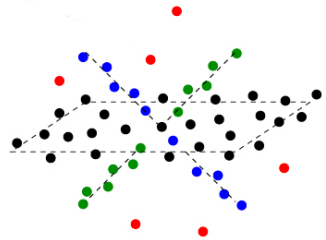
   Suppose the measurement $\tilde{\mathbf{z}} = \mathbf{z} + \mathbf{n}$, where $\mathbf{n}$ is (Gaussian) noise.
   Then $p_1(\tilde{\mathbf{z}}) \neq 0, \cdots, p_r(\tilde{\mathbf{z}}) \neq 0$.



3. Outlier issue.

   Outliers $\{\mathbf{y}\}$ may appear in the data set, leading to:

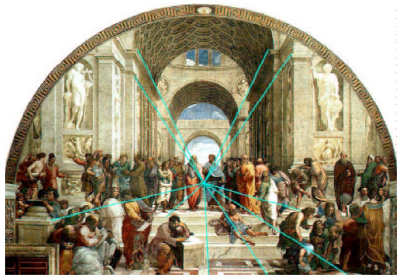
$$p_j(\mathbf{y}) \gg 0.$$



4. What to do with affine subspaces?

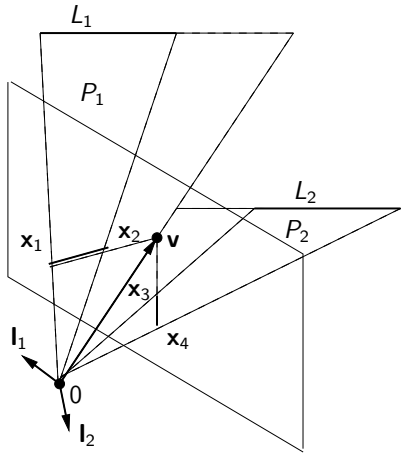
Vanishing-Point Detection:

1. Perspective projection of 2 parallel lines in space intersect at *vanishing point* in image plane.



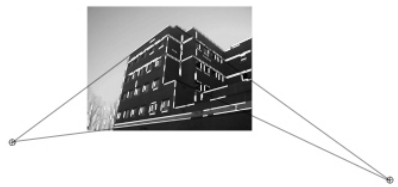
2. Geometry of a family of parallel lines:
   - The co-image of a line L_1 is $\mathbf{l}_1 = \mathbf{x}_1 \times \mathbf{x}_2$.
   - Given pre-images of two parallel lines L_1 and L_2, $(0, \mathbf{v})$ is on the intersection of P_1 and P_2.

   $\Rightarrow$ Any co-image $\mathbf{l}_i$ of a line parallel to L_1 satisfies: $\mathbf{l}_i \perp \mathbf{v}$.



3. Multiple vanishing points:
   - Multiple families of parallel lines correspond to multiple vanishing points $\mathbf{v}_1, \ldots, \mathbf{v}_n$.
   - Any co-image of a line in the families must satisfy

$$(\mathbf{l}^T \mathbf{v}_1)(\mathbf{l}^T \mathbf{v}_2) \cdots (\mathbf{l}^T \mathbf{v}_n) = 0.$$



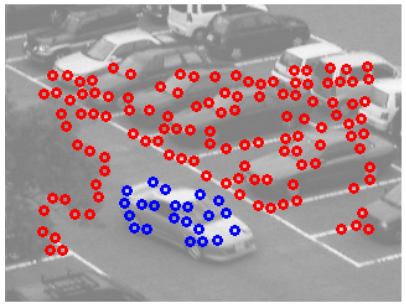
$\Rightarrow$ Segmenting parallel line families is equivalent to segmenting 2-D subspaces in $\mathbb{R}^3$.

Motion Segmentation under 3-D Affine Projection

- Affine projection of a single rigid body:

1. Object features $\mathbf{p}_1, \ldots, \mathbf{p}_N \in \mathbb{R}^3$ are tracked in F frames.

parking-lot movie



2. Denote $\mathbf{m}_{ij} \in \mathbb{R}^2$ as the image under 3-D affine projection:

$$\mathbf{m}_{ij} = A_j \mathbf{p}_i + \mathbf{b}_j \in \mathbb{R}^2, \quad i = 1, \ldots, N; j = 1, \ldots, F.$$

3. For each p_i,

$$\mathbf{z}_i = \begin{bmatrix} \mathbf{m}_{i1} \\ \vdots \\ \mathbf{m}_{iF} \end{bmatrix} \in \mathbb{R}^{2F}, \quad i = 1, \ldots, N.$$

- Multibody segmentation:
Given K independent objects, segment $\mathbf{z}_1, \ldots, \mathbf{z}_N$ that belong to different motions.

$$\Rightarrow \begin{Bmatrix} (A_{1,1}, \mathbf{b}_{1,1}) & \cdots & (A_{1,F}, \mathbf{b}_{1,F}) \\ \vdots & \vdots & \vdots \\ (A_{K,1}, \mathbf{b}_{K,1}) & \cdots & (A_{K,F}, \mathbf{b}_{K,F}) \end{Bmatrix}.$$

Suppose $\mathbf{p}_1, \ldots, \mathbf{p}_N \in \mathbb{R}^3$ are from a single object:

- Stack corresponding images in F frames: $\mathbf{z}_i = \begin{bmatrix} \mathbf{m}_{i1} \\ \mathbf{m}_{i2} \\ \vdots \\ \mathbf{m}_{iF} \end{bmatrix} \in \mathbb{R}^{2F}$, $i = 1, \cdots, N$:

$$W \doteq [\mathbf{z}_1 \cdots \mathbf{z}_N]_{2F \times N} = \begin{bmatrix} A_1 & \mathbf{b}_1 \\ \vdots & \vdots \\ A_F & \mathbf{b}_F \end{bmatrix}_{2F \times 4} \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_N \\ 1 & \cdots & 1 \end{bmatrix}_{4 \times N}.$$

  $\Rightarrow \mathbf{z}_i \in \mathbb{R}^{2F}$ lives in a subspace of dimension 4.

- When all $\mathbf{p}_i$'s are coplanar, there exists a world coordinate system such that $\mathbf{p}_i = \begin{bmatrix} x_i \\ y_i \\ 0 \end{bmatrix}$.

$$W \doteq [\mathbf{z}_1 \cdots \mathbf{z}_N]_{2F \times N} = \begin{bmatrix} A_1 & \mathbf{b}_1 \\ \vdots & \vdots \\ A_F & \mathbf{b}_F \end{bmatrix}_{2F \times 4} \begin{bmatrix} x_1 & \cdots & x_N \\ y_1 & \cdots & y_N \\ 0 & \cdots & 0 \\ 1 & \cdots & 1 \end{bmatrix}_{4 \times N}.$$

  $\Rightarrow \mathbf{z}_i \in \mathbb{R}^{2F}$ lives in a subspace of dimension 3.

Multiple rigid bodies under affine projection

Segmenting multiple rigid bodies under affine camera projection is equivalent to segmenting multiple subspaces of dimension 3 or 4.

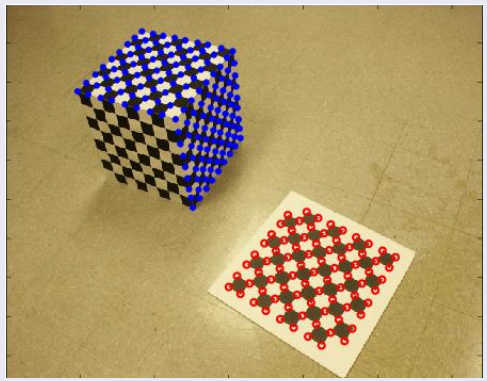
Perspective projections (bilinear constraints)

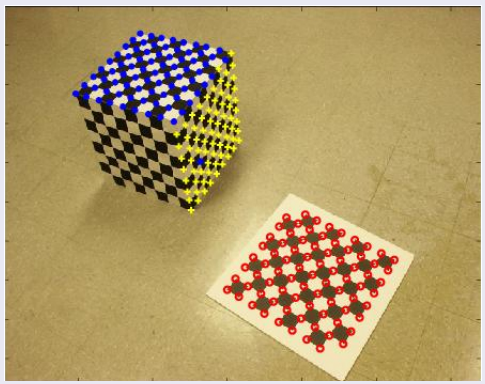
- Consider *one* epipolar constraint:
$$p(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_2^T F \mathbf{x}_1 = 0,$$
which is a bilinear constraint on $(\mathbf{x}_1, \mathbf{x}_2)$.
- For N epipolar constraints,
$$(\mathbf{x}_2^T F_1 \mathbf{x}_1)(\mathbf{x}_2^T F_2 \mathbf{x}_1) \cdots (\mathbf{x}_2^T F_N \mathbf{x}_1) = 0.$$
- Similar techniques exist to convert bilinear constraints to a subspace-segmentation problem:
$$\mathbf{x}_2^T F_i \mathbf{x}_1 \Leftrightarrow (\mathbf{x}_1 \otimes \mathbf{x}_2)^T F_i^s = 0.$$

Multiple rigid bodies under perspective projection

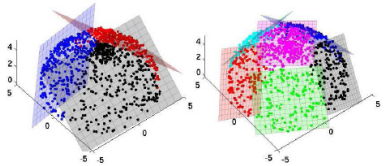
Segmenting multiple rigid bodies under perspective projection (bilinear constraints) can be converted to a subspace segmentation problem.

Be careful about degenerate structures:

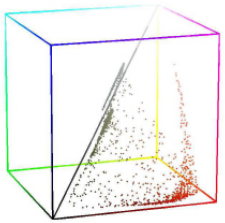




Approximate nonlinear structures



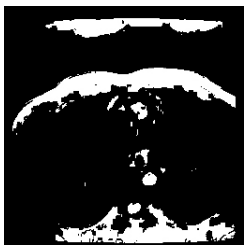
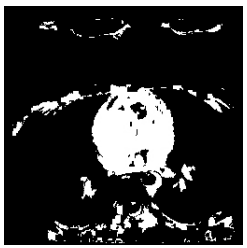
- Processing image and video data: Critical to handle large non-Gaussian noise by nonlinear structures.



- Image Segmentation

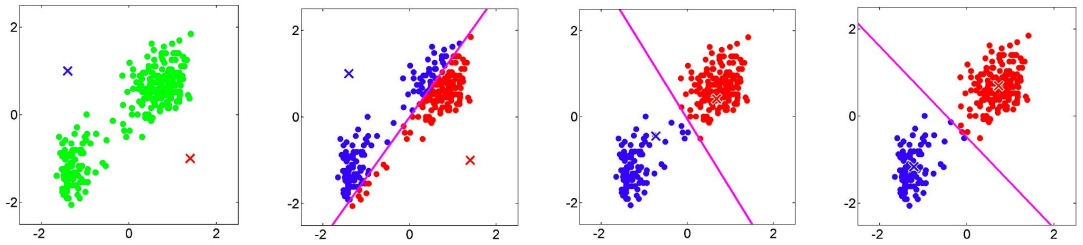
- Dynamic Texture Segmentation

Heart

K-Means and K-Subspaces

- K-Means for segmenting K Gaussian clusters:

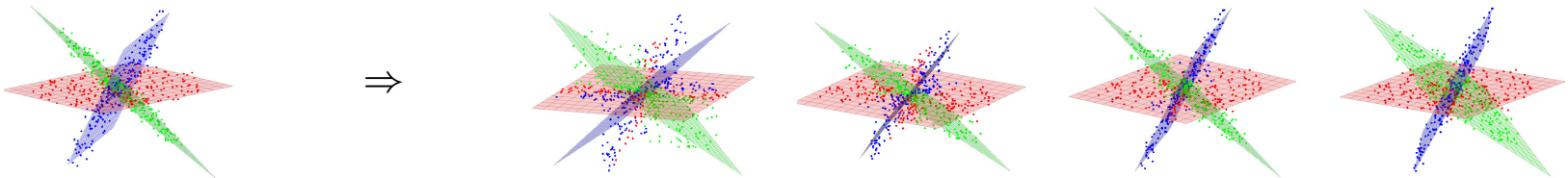


- K-Subspaces for subspace arrangements [Ho et al., 2003]:

  1. Initialization: Set initial values of orthogonal matrices $\hat{U}_i^{(0)} \in \mathbb{R}^{D \times d_i}$ for $i = 1, \ldots, N$. Let $m = 0$.
  2. Segmentation: For each sample $\mathbf{z}_k$, assign it to group $\hat{X}_i^{(m)}$ if

$$i = \arg\min_l \|\mathbf{z}_k - \hat{U}_l^{(m)} (\hat{U}_l^{(m)})^T \mathbf{z}_k\|^2.$$

  3. Estimation: Apply PCA to each subset $\hat{X}_i^{(m)}$ and obtain new estimates for the subspace bases $\hat{U}_i^{(m+1)}$.
  4. Let $m \leftarrow m + 1$, and repeat step 2 and 3 until the segmentation does not change.



- Question: What is your estimate of where the local minima are?

Expectation-Maximization

- Setup:
  1. Given $\{\mathbf{z}_1, \cdots, \mathbf{z}_N\}$, amend a latent membership for each sample $(\mathbf{z}_i, \eta_i) \in \mathbb{R}^D \times \mathbb{N}$ and a probability $\pi_{i,j} \doteq p(\eta_i = j)$.
  2. Assume a Gaussian noise model for each subspace model $\theta_i = (B_i, \sigma_i, \pi_i)$:

$$p(\mathbf{z}|\eta = i) \doteq \frac{1}{(2\pi)^{(D-d_i)/2}\sigma_i} \exp\left(-\frac{\mathbf{z}^T B_i B_i^T \mathbf{z}}{2\sigma_i^2}\right),$$

  where $B_i \in \mathbb{R}^{D\times(D-d_i)}$ is a orthogonal matrix.

- Algorithm:
  1. Initialization: Set initial values for $\hat{\theta}_i^{(0)} = \{\hat{B}_i^{(0)}, \hat{\sigma}_i^{(0)}, \hat{\pi}_i^{(0)}\}$ for $i = 1, \ldots, n$. Set $m = 0$.
  2. Expectation: Compute the expected value of w_{ik} as

$$w_{ik}^{(m)} \doteq p(\eta_k = i|\mathbf{z}_k, \hat{\theta}^{(m)}) = \frac{\hat{\pi}_i^{(m)} p(\mathbf{z}_k|\eta_k = i, \hat{\theta}^{(m)})}{\sum_{l=1}^N \hat{\pi}_l^{(m)} p(\mathbf{z}_k|\eta_k = l, \hat{\theta}^{(m)})}. \tag{1}$$

  3. Maximization: Using the expected values $w_{ij}^{(m)}$, compute $\hat{\theta}^{(m+1)}$.

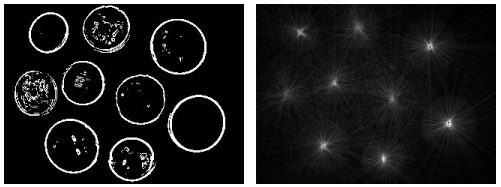
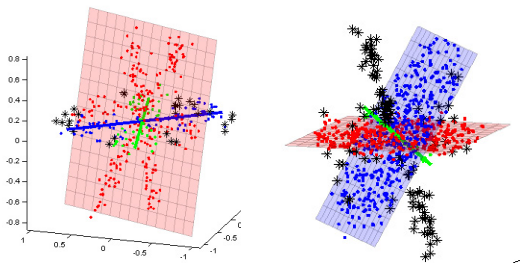
$$\begin{aligned} \hat{B}_i^{(m+1)} &= \text{the eigenvectors associated with the smallest } D - d_i \text{ eigenvalues of} \\ &\quad \text{the matrix } \textstyle\sum_{k=1}^n w_{ik}^{(m)} \mathbf{z}_k \mathbf{z}_k^T. \\ \hat{\pi}_i^{(m+1)} &= \frac{\sum_{k=1}^n w_{ik}^{(m)}}{n}. \\ \left(\hat{\sigma}_i^{(m+1)}\right)^2 &= \frac{\sum_{k=1}^n w_{ik}^{(m)} \|(\hat{B}_i^{(m+1)})^T \mathbf{z}_k\|^2}{(D-d_i)\sum_{k=1}^n w_{ik}^{(m)}}. \end{aligned} \tag{2}$$

  4. Let $m \leftarrow m + 1$, and repeat step 2 and 3 until the update in the parameters is small enough.

Random Sample Consensus (RANSAC)

- Consensus-based algorithms:

Hough: [Ballard 1981, Lowe 1999]
RANSAC: [Fischler & Bolles 1981, Torr 1997]
Least Median Estimate (LME): [Rousseeuw 1984, Steward 1999]



- Estimating subspace arrangements via RANSAC
  1. RANSAC-on-Union: Simultaneously estimation multiple subspaces.
     Complexity: Subspaces $(2, 2, 2)$ with samples $(200, 200, 200)$. Probability for a good sample set:
     $$\frac{(200 \cdot 199)^3 3!}{600 \cdot 599 \cdots 595} = 0.8\%.$$
  2. RANSAC-on-Subspaces: Estimate one subspace at a time.
     Complexity: Subspaces $(2, 2, 2)$ with samples $(200, 200, 200)$. Probability for a good sample set:
     $$\frac{3(200 \cdot 199)}{600 \cdot 599} = 33\%.$$
- RANSAC-on-Subspaces with subspace degeneracy detection
  1. Start with the highest-dimensional model. Find a sample subset with highest consensus.
  2. Verify if the subset can be fitted with lower-dimensional models. If so, temporarily discard the samples, and re-estimate the model from the remaining samples.