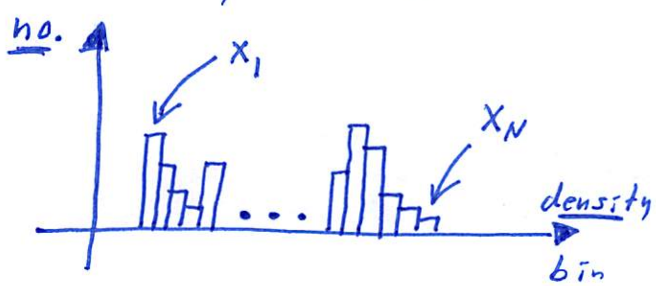
Bernd Hamann
Stratovan

9-19-2016

■ On Solving Under-determined Linear Systems & Material/Object Classification and Detection

① A material is understood as its associated "characteristic histogram" – e.g., a discrete density histogram ('binned' and using N bins):



→ Histogram interpreted as an N-dim positional vector – that is normalized with only non-negative components / coordinates:

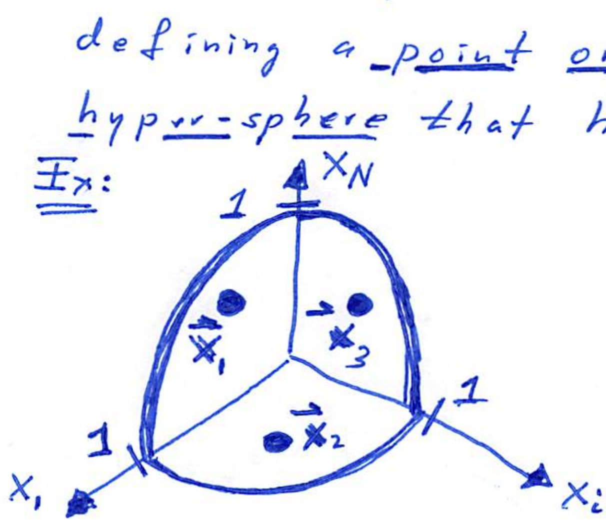
$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

$$\Rightarrow \|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\sum_{i=1}^{N} x_i^2} = \ell$$

$\Rightarrow$ normalized vector: $\vec{x}^{\text{norm}} = \frac{1}{\ell} \cdot \vec{x}$

② Each "pure", proto-typical material has an associated (positional) normalized vector defining a point on that part of a unit hyper-sphere that has only non-negative coordinates:

Ex:



→ $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$ are 3 points on the N-dim. hypersphere representing 3 distinct, pure proto-type materials.

Solving over- and under-determined sets of equations

Suppose

$$\boxed{\mathbf{y} = M\mathbf{x}}$$

where M is a $n \times m$ matrix, $\mathbf{y}$ is a known n-vector and $\mathbf{x}$ is an unknown m-vector.

1 First, assume $n > m$. In this case there are more constraints than unknowns, and the system is overdetermined, with no solutions (except for degenerate cases). We can find a least-squares solution that minimizes the error $(\mathbf{y} - M\mathbf{x})$. We want to find $\mathbf{x}$ that minimizes OVER-DETERMINED

$$\|\mathbf{y} - M\mathbf{x}\|^2$$

or

$$(\mathbf{y} - M\mathbf{x})^T(\mathbf{y} - M\mathbf{x})$$

or

$$\mathbf{y}^T\mathbf{y} - \mathbf{y}^T M\mathbf{x} - \mathbf{x}^T M^T\mathbf{y} + \mathbf{x}^T M^T M\mathbf{x}$$

Differentiating w.r.t. $\mathbf{x}$ and setting the result equal to zero yields

$$-(\mathbf{y}^T M)^T - (M^T\mathbf{y}) + 2M^T M\mathbf{x} = 0$$

so

$$\boxed{\mathbf{x} = (M^T M)^{-1} M^T\mathbf{y}} \iff M^T M\mathbf{x} = M^T\mathbf{y}$$

where $(M^T M)^{-1} M^T$ (a $m \times n$ matrix) is called a pseudo-inverse.

2 Second, assume $n < m$. In this case there are fewer constraints than unknowns, and the system is underdetermined, with an infinite number of solutions. We can pick one of these solutions by finding the smallest one. That is, we will (*) minimize $\mathbf{x}$ subject to the constraint $\mathbf{y} = M\mathbf{x}$. The method of Lagrange multipliers has us add a term to the quantity to be minimized: UNDER-DETERMINED !

$$\|\mathbf{x}\|^2 + \boldsymbol{\lambda}^T(\mathbf{y} - M\mathbf{x})$$

(*) This is just ONE possibility to define a unique solution! QUESTION: "What is the 'best way' to define a unique solution for our purpose — classification ???"

Differentiting w.r.t $\mathbf{x}$ and setting the result equal to zero yields

$$2\mathbf{x} - M^T\boldsymbol{\lambda} = 0$$

We can't just solve for $\boldsymbol{\lambda}$ since M is not a square matrix, but we can premultiply by M to obtain

$$2M\mathbf{x} - MM^T\boldsymbol{\lambda} = 0$$

and using $\mathbf{y} = M\mathbf{x}$ gives us

$$2\mathbf{y} = MM^T\boldsymbol{\lambda}$$

so

$$\boldsymbol{\lambda} = 2(MM^T)^{-1}\mathbf{y}$$

and hence

$$\boxed{\mathbf{x} = M^T(MM^T)^{-1}\mathbf{y}} \quad (*) \quad \text{(here: } p = 0\text{)}$$

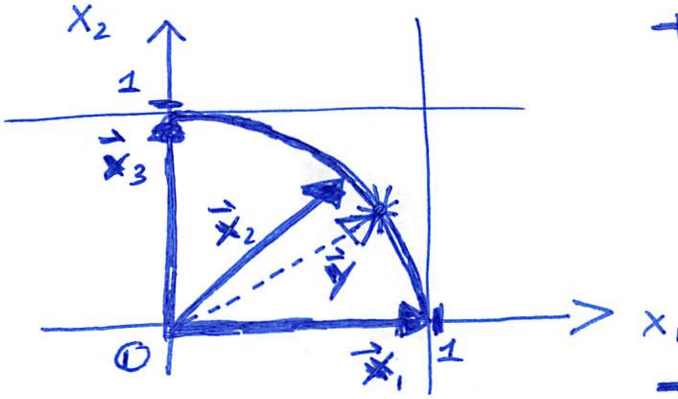
where $M^T(MM^T)$ (a $m \times n$ matrix) is called a pseudo-inverse.

(*) More general: FIND $\mathbf{x}$ such that $\|\mathbf{x} - p\|_2$ is minimal.
"Solution $\mathbf{x}$ closest to a specified p"

$$\boxed{\mathbf{x} = p + M^T(MM^T)^{-1}\cdot(\mathbf{y} - Mp)}$$

REF: Boehm & Prautzsch, Numerical Methods, p. 41.

③ **Example**: Under-determined problem in the plane — Using **3** **2D** **vectors**



→ 3 normalized vectors with only positive x_1 and x_2 coordinates:

$\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{x}_2 = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}$, $\vec{x}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

→ $\vec{y} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to be represented by "the best" linear combination of $\vec{x}_1, \vec{x}_2, \vec{x}_3$:

→ Determining the solution that minimizes $\vec{x} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ in $M\vec{x} = \vec{y}$ leads to the system

$$\alpha \vec{x}_1 + \beta \vec{x}_2 + \gamma \vec{x}_3 = \vec{y}$$

$$\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & \sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Leftrightarrow M\vec{\alpha} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\rightarrow \underline{\vec{\alpha}} = M^T (M M^T)^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3/4 & -1/4 \\ -1/4 & 3/4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 3/4 & -1/4 \\ \sqrt{2}/4 & \sqrt{2}/4 \\ -1/4 & 3/4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underline{\frac{1}{4} \begin{pmatrix} 3 & -1 \\ \sqrt{2} & \sqrt{2} \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}$$

Bernd Hamann
Stratovan

9-19-2016

(Example - cont'd.)

→ Let's determine "the optimal" coefficient vectors for $\vec{y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\vec{y} = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}$:

(i) $\vec{y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

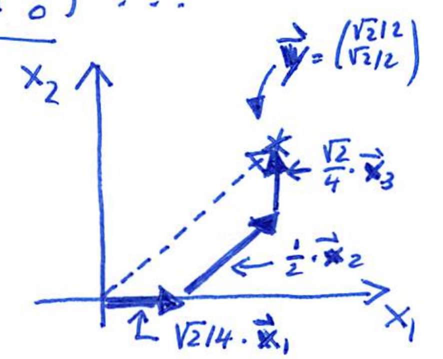
$$\Rightarrow \vec{\alpha} = \frac{1}{4} \begin{pmatrix} 3 & -1 \\ \sqrt{2} & \sqrt{2} \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 3 \\ \sqrt{2} \\ -1 \end{pmatrix}$$

⇒ • IS "COORDINATE" VALUE -1 ALLOWABLE (for our application)?

• WHILE THIS $\vec{\alpha}$ COEFFICIENT VECTOR IS 'OPTIMAL' (in terms of its length), IS IT "BETTER" THAN $\vec{\alpha} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$???

(ii) $\vec{y} = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$\Rightarrow \vec{\alpha} = \frac{1}{4} \begin{pmatrix} 3 & -1 \\ \sqrt{2} & \sqrt{2} \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} \sqrt{2} \\ 2 \\ \sqrt{2} \end{pmatrix} \rightarrow$$

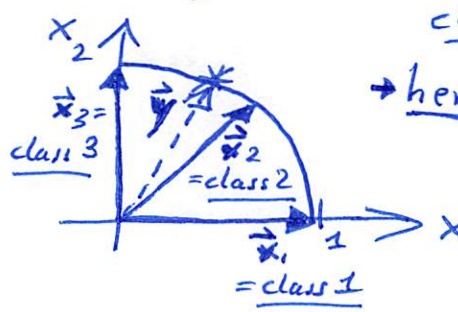


⇒ • IS THIS 'OPTIMAL' COEFFICIENT VECTOR "BETTER" THAN $\vec{\alpha} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$???

④ Relationship to Classification – "Correct Classification"

"Our Context:

$\vec{x}_1, \vec{x}_2, \vec{x}_3$ are 3 "material classes"



→ here: $\vec{y}$ is close to $\vec{x}_2$

→ Is this the "proper question" to be answered?

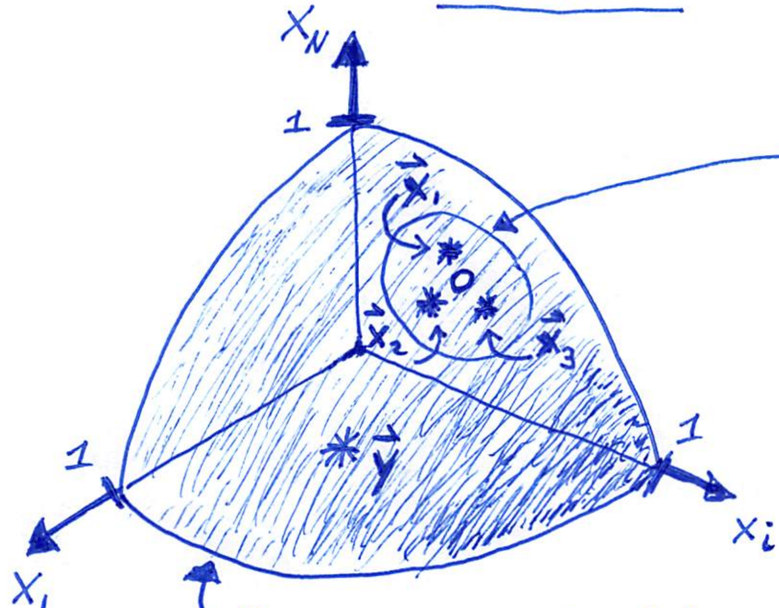
• Given "material $\vec{y}$", to which class does it belong, i.e., is $\vec{y}$ close to an $\vec{x}_i$ such that it can be viewed as belonging to $\vec{x}_i$?

Bernd Haman
Stratovan

9- 26- 2016

⑤ Possibility to 'avoid' using a linear systems approach – Using a SCALAR PRODUCT / ANGLE criterion instead

→ Assumption: All histogram data consists of N-dim. points/vectors with only non-negative coordinates – and data is normalized!



Cluster C_j, $j = 1 \ldots L$, consisting of K_j (here: $K_j = 3$) associated points $\vec{x}_1, \vec{x}_2, \vec{x}_3$ – "defining" one CLASS/ MATERIAL TYPE □

"positive part" of hyper-sphere $x_1^2 + \ldots + x_N^2 = 1$, $0 \leq x_i \leq 1$, $i = 1 \ldots N$

• PROBLEM:
Given new point $\vec{y}$, does it belong to any of the clusters/classes?

→ (i) One can associate a "center $\vec{c}_j$" with each cluster C_j, by averaging its associated points $\vec{x}_1, \vec{x}_2, \ldots$ (for example):

$$\vec{c}_j = \frac{1}{K_j} \sum_{l=1}^{K_j} \vec{x}_l \Big/ \left\| \frac{1}{K_j} \sum_{l=1}^{K_j} \vec{x}_l \right\| \quad (\text{'o' in figure})$$

$\vec{a} \cdot \vec{b}$ = scalar product of $\vec{a}$ and $\vec{b}$

(ii) DOES $\vec{y}$ REPRESENT A MATERIAL IN C_j?
$\Longleftrightarrow$ $\vec{y} \cdot \vec{c}_j$ CLOSE TO 1.0? YES $\Rightarrow$ $\vec{y}$ OF CLASS C_j

(iii) HOW "CLOSE" IS $\vec{y}$ TO MATERIAL $\vec{x}_l$? CONSIDER $\vec{y} \cdot \vec{x}_l$ ■

BH