The Geometry of $\mathbb{R}^n$

Orthogonal Coordinate Systems and Least Squares Problems

February 9, 2010

Scott Strong

sstrong@mines.edu

Colorado School of Mines

Overview/Keywords/References

Advanced Engineering Mathematics **Slide Set Two**

Coordinate Systems, Inner-product, Orthogonality, Normal Equations

Reference Text: N/A **Example: Least-Squares Problems**

Before We Begin

Quote of Slide Set Two
Not that the propositions of geometry are only approximately true, but that they remain absolutely true in regard to that Euclidean space which has been so long regarded as being the physical space of our experience.
Arthur Cayley : Sadleirian Professor (1863-1895)

A Little Back Story

If $\mathbf{A} \in \mathbb{R}^{n\times n}$ is an invertible matrix then every $\mathbf{b}_{n\times 1} \in$Col$(\mathbf{A})$ and there exist unique constants $x_1, x_2, x_3, \ldots, x_n$ such that,

$$\sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b}. \tag{1}$$

If we wanted to find these unknown scalars then we would apply row-reduction to $[\mathbf{A}|\mathbf{b}]$.
<u>Question</u>: What is **b** really?
<u>Answer</u>: A linear combination of the standard-basis vectors.

$$\mathbf{b} = \mathbf{I}\mathbf{b} = \sum_{i=1}^{n} b_i \hat{\mathbf{e}}_i, \text{ where } \left[\hat{\mathbf{e}}_i\right]_j = \delta_{ij} \tag{2}$$

We say that $(b_1, b_2, b_3, \ldots, b_n)$ are the coordinates of **b** relative to the standard-basis.

Coordinate Systems and Invertible Matrices

If **A** is an invertible matrix then the columns of **A** form another basis for $\mathbb{R}^n$ and the linear combination,

$$\sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b}, \qquad (3)$$

expresses **b** in this basis. We say that $(x_1, x_2, x_3, \ldots, x_n)$ are the coordinates of **b** relative to the basis $B_{\mathbb{R}^n} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_n\}$ and

$$\underbrace{\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}}_{\text{Coord. of } \mathbf{b} \text{ relative to } B_{\mathbb{R}^n}} \iff \underbrace{\mathbf{b} = \mathbf{A}\mathbf{x}}_{\mathbf{b} \text{ relative to standard-basis}}$$

Key Point: Every $n \times n$ invertible matrix defines an alternate coordinate system for $\mathbb{R}^n$ and matrix multiplication defines a change between these coordinate systems.

Inner-Product

When thinking about the geometry of coordinate systems it is natural to ask the question,

- How are each of the basis vectors separated relative to one another?

To address this question we recall that for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $n = 2, 3$ we have the scalar/dot-product,

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^{\mathsf{T}}\mathbf{y} = [x_1 \ x_2 \ x_3] \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \sum_{i=1}^{3} x_i y_i, \tag{4}$$

where $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos(\theta) \Rightarrow \mathbf{x} \cdot \mathbf{y} = 0 \iff \theta = (2k-1)\pi/2$.
Question: Does this definition hold for the space $\mathbb{R}^n$?
Answer: Maybe. If so, then we'll call it an inner-product.

Norms/Lengths Generated by Inner-Products

Abstracting from vector calculus we have that for $n = 2, 3$,

$$||\mathbf{x}|| = |\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\mathbf{x}^{\mathsf{T}}\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i x_i} \geq 0, \tag{5}$$

which clearly holds for $n > 2$. An inner-product can always be used to generate a norm and again we ask,

- Can $\mathbf{x}^{\mathsf{T}}\mathbf{y} = ||\mathbf{x}||\,||\mathbf{y}||\cos(\theta)$ be used in general for $\mathbb{R}^n$?

Yes, if we assume the Cauchy-Schwarz inequality,

$$\left|\mathbf{x}^{\mathsf{T}}\mathbf{y}\right| \leq ||\mathbf{x}||\,||\mathbf{y}||. \tag{6}$$

Key Point: We can abstract the standard results for angle and length in $\mathbb{R}^2$ to $\mathbb{R}^n$.

Orthogonal Coordinate Systems

Now that we have length and angle back on our side we state:

- Bases are good. Orthogonal bases are better. Orthonormal bases are best.

If we assume that the vectors from $B_{\mathbb{R}^n}$ are such that $\mathbf{a}_j^\mathsf{T}\mathbf{a}_i = \delta_{ji}$ then we have,

$$\sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b} \iff \mathbf{a}_j^\mathsf{T} \sum_{i=1}^{n} x_i \mathbf{a}_i = \sum_{i=1}^{n} x_i \mathbf{a}_j^\mathsf{T}\mathbf{a}_i = \sum_{i=1}^{n} x_i \delta_{ji} = \mathbf{a}_j^\mathsf{T}\mathbf{b}, \tag{7}$$

which gives, $x_j = \mathbf{a}_j^\mathsf{T}\mathbf{b}$, for $j = 1, 2, 3, \ldots, n$.
Key Point: The problem of row-reduction has been replaced with n-many inner-product calculations.

Projections

Question: What is the geometric meaning of $x_j = \mathbf{a}_j^\top \mathbf{b}$?
Answer: An inner-product is an example of a projection.
Recall from calculus the scalar projection of **y** in the direction of **x**,

$$\text{comp}_{\mathbf{x}}\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{x}}{|\mathbf{x}|} \in \mathbb{R}. \tag{8}$$

For $x_j = \mathbf{a}_j^\top \mathbf{b}$ we say that:

- The vector **b** is projected onto the j^{th} basis vector and x_j is the signed-length of this projection onto $\mathbf{a}_j$.
- $x_j = \text{comp}_{\mathbf{a}_j}\mathbf{b} = \dfrac{\mathbf{a}_j^\top \mathbf{b}}{|\mathbf{a}_j|} = \mathbf{a}_j^\top \mathbf{b}$
- x_j is the distance from the origin **b** goes in the $\mathbf{a}_j$ direction.

Least-Squares Approximations

Question: Given $\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{m\times 1}$. What if there is no solution to this system?

Answer: Consider the affect of $\mathbf{A}^\top\mathbf{b} = \tilde{\mathbf{b}}$,

$$\left[\tilde{\mathbf{b}}\right]_i = \left[\mathbf{A}^\top\mathbf{b}\right]_i = \mathbf{a}_i \cdot \mathbf{b}, \tag{9}$$

which defines a vector whose i^{th} component is the projection of $\mathbf{b}$ in the direction of the i^{th} column of $\mathbf{A}$. Thus, $\tilde{\mathbf{b}} \in \text{Col}(\mathbf{A})$ and the system of normal equations,

$$\mathbf{A}^\top\mathbf{A}\mathbf{x} = \tilde{\mathbf{b}} = \mathbf{A}^\top\mathbf{b}, \tag{10}$$

must have a solution!

Key Point: The associated least-squares problem defined by (10) for $\mathbf{A}\mathbf{x} = \mathbf{b}$ will always have a solution that defines the coordinates of the element in Col$(\mathbf{A})$ closest to $\mathbf{b}$.

Example : Geometry

Consider the following system of equations,

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \tag{11}$$

for which there is no solution. However, the associated LSP,

$$\mathbf{A}^\mathsf{T}\mathbf{A} = 2\mathbf{I}, \quad \mathbf{A}^\mathsf{T}\mathbf{b} = \mathbf{0}, \tag{12}$$

has the solution, $\mathbf{x} = \mathbf{0}$.

Example : Polynomial Interpolation

Consider the polynomial interpolation problem defined by set two from problem 4.2 of homework 1. That is, find the constants a_0, a_1, a_2 such that $p(t) = a_0 + a_1 t + a_2 t^2$ passes through $S_2 = \{(1,12), (1,15), (3,16)\}$. This defines the system,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 12 \\ 15 \\ 16 \end{bmatrix}, \tag{13}$$

which has no solution. Clearly, no function can pass through both the first and second points in S but the associated LSP has infinitely-many solutions which define the polynomials,

$$p(t) = \left(\frac{49}{4} + 3a_2\right) + \left(\frac{5}{4} - 4a_2\right) t + a_2 t^2, \ \ a_2 \in \mathbb{R} \tag{14}$$