

ARE211, Fall 2009

LINALGEBRA4: TUE, OCT 13, 2009

PRINTED: AUGUST 27, 2009

(LEC# 14)

CONTENTS

3. Linear Algebra (cont)	1
3.11. The “graph” of a linear function from \mathbb{R}^2 to \mathbb{R}^2 (cont)	1
3.12. Computational details	3
3.13. Determinants, Rank and volume	4
3.14. Solving linear equation systems and Cramer’s Rule	6

3. LINEAR ALGEBRA (CONT)

3.11. The “graph” of a linear function from \mathbb{R}^2 to \mathbb{R}^2 (cont)

So far, we have been considering only *symmetric* matrices. Reason is that we were talking about eigenvalues, eigenvectors and definiteness. For symmetric matrices, these relationships are very clearcut. The waters are substantially muddied with nonsymmetric matrices. For our purposes, we are only interested definiteness of matrices when the matrices are Hessians, and these matrices are always symmetric, i.e., the cross-partials are identical. So the restriction is harmless. Specifically, the following facts are true:

- For any *symmetric* $n \times n$ matrix, there exists a set of eigenvectors which each have unit length and are pairwise orthogonal, i.e., for any two elements in the set, \mathbf{v}^1 and \mathbf{v}^2 , the inner product $\mathbf{v}^1 \cdot \mathbf{v}^2$ is zero.

- For any *symmetric* $n \times n$ matrix, the matrix is *positive (negative)* definite if and only if all of its eigenvalues are positive (negative).
- For any *symmetric* $n \times n$ matrix, the rank of the matrix is equal to the number of nonzero eigenvalues. More precisely, let A be a symmetric $n \times n$ matrix, let $V = \{v^1, \dots, v^n\}$ be a set of pair-wise orthogonal eigenvectors for A and let $\Lambda = \{\lambda_1, \dots, \lambda_n\}$ be a set of corresponding eigenvalues, i.e., for each i , $Av^i = \lambda_i v^i$. Then the rank of A is equal to the number of nonzero elements of Λ .

Semi-definite matrices

Take the following symmetric matrix: $D = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Note that both vectors are pointing in the same direction, so that we know it's singular, and that we are going to get something degenerate

- Indeed, look at what happens to the unit circle: any vector (x_1, x_2) gets mapped into $(x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (because the second vector in the matrix is in fact $2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$), i.e., the unit circle collapse to a line which is of course the subspace of \mathbb{R}^2 that is spanned by the columns of D
- Note that it does indeed have two orthogonal eigenvectors: $v^1 = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ and $v^2 = \left(\frac{-2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$.
- v^2 gets mapped to $(0,0)$, i.e., (v_1^2, v_2^2) gets mapped into $(v_1^2 + 2v_2^2) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (-2 + 2 * 1) \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0$
- Corresponding eigenvalues are: $(5,0)$
- The first eigenvector gets pulled out along the line
- The second goes nowhere

Similarly, a 3×3 symmetric matrix which is semidefinite will take the unit sphere either to an ellipse that lives inside a 2-dimensional subspace of \mathbb{R}^3 (if it has rank 2) or else simply a line in \mathbb{R}^3 (if it has rank 1).

The set of all vectors that are mapped by an $n \times n$ matrix A to zero is called the *nullspace* or *kernel* of A . Note that it is a vector subspace of \mathbb{R}^n . Properties:

- the dimension of the null-space is equal to the number of zero eigenvalues
- the null-space of A is spanned by the eigenvectors whose corresponding eigenvalues are zero.
- Obvious but fundamental theorem: the dimension of the null-space + the rank of $A = n$

What's the null-space of a matrix with full rank? Obviously the zero-dimensional subspace $\{0\}$.

Non-symmetric matrices.

There are nonsymmetric matrices that have no (real) eigenvectors or eigenvalues. One example is $E = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ which swivels every vector thru 90 degrees. It is positive semidefinite, but not positive definite *even though it has full rank*. (By the way, it is also negative semidefinite but not negative definite.) This illustrates that the relationship between definiteness and full rank that we've talked about holds *only* for symmetric matrices.

Here's yet another matrix that is positive definite, but is qualitatively the same as the preceding one. $E = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Everything gets rotated clockwise through 45 degrees.

Non-square matrices. Why can't you have eigenvalues of a nonsquare matrix? Answer: for a nonsquare matrix, domain and range are different, so you can't have point in the image being a scalar multiple of the point in the domain

3.12. Computational details

You should be able to compute the following,

- Eigenvectors, at least for a 2x2 system
- Eigenvalues, at least for a 2x2 system
- Determinant for a 2x2 and for a 3x3

- Inverse for a 2×2 and for a 3×3
- The solution to a 2×2 or 3×3 equation system

3.13. Determinants, Rank and volume

An important interpretation of the determinant is its relationship to *area* or *volume*. Two versions of this relationship:

(1) Interpretation relating to the columns of the matrix

- The determinant of a 2×2 (3×3) matrix is equal to the area (volume) of the parallelogram (parallelepiped) defined by its columns.
- For a 2×2 , notice that if the columns are (nearly) collinear, determinant is (nearly) zero
- For a 3×3 , notice that if the columns are (nearly) in a plane, determinant is (nearly) zero

(2) Interpretation relating to the eigenvectors of the matrix

- For a 2×2 matrix A , the absolute value of the determinant of A is equal to the ratio of the *area* of the image of the unit circle under A to the area of the unit circle itself.
 - Area of any circle is πr^2 .
 - Area of the ellipse is $\pi \prod_{i=1}^2 r_i$ where the r_i 's are the radii of the axes of the ellipse, i.e., the length from the center of the ellipse to the tip of each axes.
 - Area of the *unit* circle is π .
 - Area of the image of the unit circle is $\pi \prod_{i=1}^2 (|\lambda_i| \|v^i\|)$ where the v^i 's are the eigenvectors. But the eigenvectors in the unit circle have unit length, so that the area of the image of the unit circle is $\pi \prod_{i=1}^2 |\lambda_i|$.
- For a 3×3 matrix A , the absolute value of the determinant of A is equal to the ratio of the *volume* of the image of the unit sphere under A to the volume of the unit sphere itself.
 - Volume of any sphere is $\frac{4}{3}\pi r^3$.

- Volume of the ellipsoid? is $\frac{4}{3}\pi \prod_{i=1}^3 r_i$ where the r_i 's are the radii of the axes of the ellipsoid, i.e., the length from the center of the ellipse to the tip of each axes.
- Volume of the unit sphere is $\frac{4}{3}\pi$.
- Area of the image of the unit sphere is $\frac{4}{3}\pi \prod_{i=1}^3 (|\lambda_i| |v^i|)$ where the v^i 's are the eigenvectors. But the eigenvectors in the unit circle have unit length, so that the volume of the image of the unit sphere is $\frac{4}{3}\pi \prod_{i=1}^3 |\lambda_i|$.
- Take matrix A : swivel \mathbf{v}^1 and \mathbf{v}^2 closer together till they are almost pointing in the same direction. Observe that the image of the unit circle becomes a narrower and narrower ellipse until in the limit it collapses to a line.
- Relationship between eigenvalues and the rank of a symmetric matrix: the rank of a (symmetric) matrix is equal to the number of nonzero eigenvalues of the matrix:
 - note that the rank of a 2×2 matrix is one precisely when the unit circle gets mapped to a line;
 - note that the rank of a 3×3 matrix is two when the unit sphere gets mapped to a (flat) ellipse and one when the unit sphere gets mapped to a line.
- What about the *signs* of the determinants? We cannot determine the sign by looking just at the image of the circle. We need to consider the “orientation” of the ellipse. As noted at the end of the previous lecture, when we discussed matrices A and C , a 2×2 matrix with a positive determinant “preserves” orientation, while a 2×2 matrix with a negative determinant reverses it. That is, a 2×2 matrix A *preserves* orientation if whenever \mathbf{v}^1 lies to the left of \mathbf{v}^2 , then $A\mathbf{v}^1$ lies to the left of $A\mathbf{v}^2$ while, it *reverses* orientation if whenever \mathbf{v}^1 lies to the left of \mathbf{v}^2 , then $A\mathbf{v}^1$ lies to the right of $A\mathbf{v}^2$.
- Might it happen that a 2×2 matrix preserves orientation for one pair of vectors and reverses it for another? Answer: not if the matrix has full rank, i.e., nonzero determinant. Proof would require an intermediate value theorem argument: suppose it could happen that for one pair of non-collinear vectors A preserves orientation, while for another, it reverses it. Then by continuity and the intermediate value theorem, it must be the case that for some intermediate pair of non-collinear vectors, the matrix

neither reserves or preserves orientation, i.e., the images of these two non-collinear vectors are collinear. But this can't happen for a full-rank matrix.

- The above makes sense for a 2×2 matrix, but what about an $n \times n$ matrix. In \mathbb{R}^2 , for any pair of non-collinear vectors \mathbf{v}^1 and \mathbf{v}^2 in \mathbb{R}^2 , there are exactly two possibilities: *either* \mathbf{v}^1 lies to the left of \mathbf{v}^2 *or* \mathbf{v}^1 lies to the right of \mathbf{v}^2 . So “preserve” vs “reverse” are well-defined concepts. But for an $n \times n$ matrix, if $n > 2$, then there are infinitely many directions and so orientation seems no longer to be an either/or proposition. I've yet to understand the answer to this.

3.14. Solving linear equation systems and Cramer's Rule

Two interpretations of the solution to $A\mathbf{x} = \mathbf{b}$:

- Row interpretation: a solution \mathbf{x} is a point that lies in the intersection of the level sets defined by the rows of A and \mathbf{b} .
- Column interpretation: a solution \mathbf{x} is a vector of weights that generate \mathbf{b} as a linear combination of the columns of A .

Cramer's rule is an easy way of computing the value of \mathbf{x} , given that $A\mathbf{x} = \mathbf{b}$.

If $A\mathbf{x} = \mathbf{b}$, then $x_i = \frac{|A^i|}{|A|}$, where $|A^i|$ denotes the matrix constructed by taking the columns of A and replacing the i 'th column with the vector \mathbf{b} . (Note that this notation is not standard — don't quote me on it.)

Some intuition, using the fact that the determinant of a 2×2 matrix $A = [\mathbf{v}^1, \mathbf{v}^2]$ is a measure of the area defined by the columns of A .

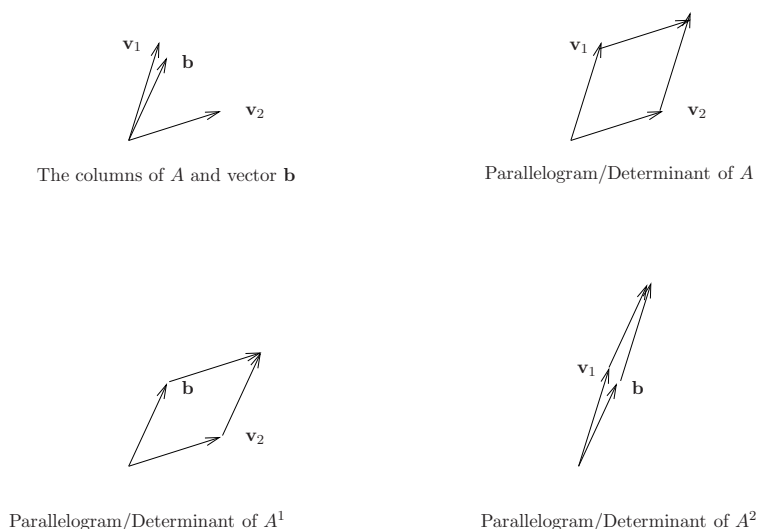


FIGURE 1. Geometric Interpretation of Cramer's Rule

- Draw the two columns of A , then draw \mathbf{b} as a convex combination of these columns, but have \mathbf{b} be much closer to \mathbf{v}^1 , the first column of A .
- Using the column interpretation of \mathbf{x} , what can we say by inspection about the relative magnitudes of x_1 and x_2 ? Ans: $x_1 \approx 1$; $x_2 \approx 0$.
- Now look at the matrix $[\mathbf{v}^1, \mathbf{b}]$; the two column vectors are very close together; while $[\mathbf{b}, \mathbf{v}^2]$ looks pretty much like the matrix A .
- Now: by Cramer's rule, $x_1 = \frac{|[\mathbf{b}, \mathbf{v}^2]|}{|A|}$; since the two matrices are virtually the same, have the same determinant, so that $x_1 \approx 1$.
- On the other hand, by Cramer's rule, $x_2 = \frac{|[\mathbf{v}^1, \mathbf{b}]|}{|A|}$; since the numerator is virtually zero, we have $x_2 \approx 0$.
- Conclude that it works!

What about the signs of the determinants? It looks like the following will create a problem. Given any matrix A made up of two column vectors $\mathbf{v}^1, \mathbf{v}^2$, and any vector \mathbf{b} that belongs to the nonnegative cone defined by them, the solution to $A\mathbf{x} = \mathbf{b}$ must be nonnegative *regardless of the order in which \mathbf{v}^1 and \mathbf{v}^2 appear in the matrix*. Specifically, for $A = [\mathbf{v}^1, \mathbf{v}^2]$ and $C = [\mathbf{v}^2, \mathbf{v}^1]$ the solutions to

$A\mathbf{x} = \mathbf{b}$ and to $C\mathbf{x} = \mathbf{b}$ must both be nonnegative. But the signs of the determinants of A and C will be opposite. So how come Cramers' rule gives the right answer for both?

The answer is that if $\det C = -\det A$ then necessarily, for each i , $\det C^i = -\det A^i$, so the two sign reversals cancel each other out. To see this, note that $\det A = \det[\mathbf{v}^1, \mathbf{v}^2]$ iff \mathbf{v}^2 is reached by moving counter-clockwise from \mathbf{v}^1 . But in this case, if \mathbf{b} is in the nonnegative cone defined by these two vectors, then \mathbf{b} will be counter-clockwise from \mathbf{v}^1 and clockwise from \mathbf{v}^2 . Thus the orientation of $[\mathbf{v}^1, \mathbf{v}^2]$ and $[\mathbf{v}^1, \mathbf{b}]$ will be the same. Similarly, Thus the orientation of $[\mathbf{v}^1, \mathbf{v}^2]$ and $[\mathbf{b}, \mathbf{v}^2]$ will be the same.