TU München
Fakultät für Informatik
PD Dr. Rudolph Triebel
John Chiotellis

Machine Learning for Computer Vision
Winter term 2016

Linear Algebra Refresher

0.1 Eigenvalues and Eigenvectors

They express the essential information contained in a **square** matrix.
Let A be an $n \times n$ matrix. A has n eigenvalues λ_i and n corresponding eigenvectors $\vec{v_i}$. What is unique about the eigenvectors of a matrix in contrast to other vectors is that the following holds:

$$A\vec{v_i} = \lambda_i \vec{v_i} \tag{1}$$

In words, this means that when we apply the linear transformation A to an eigenvector of A, we get back a multiple of the eigenvector. Geometrically speaking, all we do by applying A to $\vec{v_i}$ is to *scale* $\vec{v_i}$ by some number λ_i. This number is nothing else but the corresponding eigenvalue of $\vec{v_i}$.

To find the eigenvalues and eigenvectors of a matrix, we have to bring all parts of equation (1) to the left side:

$$(A - \lambda_i I)\vec{v_i} = 0 \tag{2}$$

where I is the $n \times n$ Identity matrix. Now, we know that for the left part to be zero, either $\vec{v_i}$ has to be the zero vector (which is not so interesting...) or the determinant of $(A - \lambda_i I)$ has to be zero:

$$\det(A - \lambda_i I) = 0 \tag{3}$$

This gives us a polynomial of λ_i of order n:

$$p(\lambda_i) = \alpha_n \lambda_i^n + \alpha_{n-1}\lambda_i^{n-1} + \ldots + \alpha_1 \lambda_i + \alpha_0 \tag{4}$$

The roots of this **characteristic polynomial** are the eigenvalues of A. Therefore the eigenvalues can be either real or complex numbers. Now we can compute the eigenvectors. For every λ_i that we found, we solve equation (2) (a system of n equations and n unknowns) and this gives us the eigenvectors $\vec{v_i}$.

Note that an eigenvalue may appear more than one time as a root of the characteristic polynomial. The **algebraic multiplicity** of an eigenvalue is the number of times it appears as a root of the characteristic polynomial (4). The **geometric multiplicity** of an eigenvalue is the number of (distinct) corresponding eigenvectors it has.

0.2 Eigendecomposition

Eigendecomposition is the rewriting of A as a product (factorization) of matrices of the eigenvectors and eigenvalues:

$$A = V\Lambda V^{-1} \tag{5}$$

where V is a matrix with the eigenvectors $\vec{v_i}$ stacked as columns:

$$V = \begin{bmatrix} \vec{v_1} & \vec{v_2} & \dots & \vec{v_n} \end{bmatrix} \tag{6}$$

and Λ is a diagonal matrix with the eigenvalues in corresponding order:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \tag{7}$$

In order for A to be "eigendecomposable", it has to be square and *diagonalizable*.

0.3 Matrix Properties Definitions

Having said something about eigenvalues and eigenvectors, it is useful to have the following definitions for matrices:

Diagonalizable A square matrix A is *diagonalizable* when the algebraic and geometric multiplicities of each eigenvalue of A coincide.

Invertible or Non-singular A square matrix A is *invertible* when it has full rank: $rank(A) = n$ or equivalently when its determinant is non-zero: $\det(A) \neq 0$.

Normal A square matrix A is *normal* if $AA^* = A^*A$, where A^* is the conjugate transpose of A.

Unitary A square matrix A is *unitary* if it is normal and also the product with the conjugate transpose gives the Identity: $AA^* = A^*A = I$. In other words the conjugate transpose is also the inverse: $A^* = A^{-1}$.

Orthogonal A square matrix A is *orthogonal* if it is unitary and real: $AA^T = A^TA = I$. (It is unfortunate that the name of these matrices is not orthonormal.)

Hermitian or Self-adjoint A square matrix A is *hermitian* if $A = A^*$.

Symmetric A square matrix A is *symmetric* if $A = A^T$, namely hermitian and real.

Skew-Hermitian A square matrix A is *skew-hermitian* if $A^* = -A$.

Skew-Symmetric A square matrix A is *skew-symmetric* if $A^T = -A$.

Let us summarize these definitions in a table:

Property	**Name** ($A \in \mathbb{C}^{n\times n}$)	**Name** ($A \in \mathbb{R}^{n\times n}$)
$\mu_A(\lambda_i) = \gamma_A(\lambda_i), \forall i$	diagonalizable	diagonalizable
$\det(A) \neq 0$	invertible	invertible
$AA^* = A^*A$	normal	normal
$AA^* = A^*A = I$	unitary	orthogonal
$A = A^*$	hermitian	symmetric
$A = -A^*$	skew-hermitian	skew-symmetric

0.4 Singular Value Decomposition

Singular Value Decomposition or SVD is the generalization of eigendecomposition to any $m \times n$ matrix. In particular we can decompose any $m \times n$ matrix A to 3 other matrices as follows:

$$A = UDV^* \tag{8}$$

where U is an $m\times m$ unitary matrix, D is an $m\times n$ rectangular diagonal matrix and V is an $n \times n$ unitary matrix. Equivalently to the eigendecomposition naming conventions, the values σ_i in the diagonal of D are called *singular values* and the column vectors of U and V are called the left- and right-*singular vectors* respectively. In contrast to eigenvalues, the singular values can only be real and non-negative.
A connection to the eigenvalues can be made by noticing that the singular values of the $m \times n$ matrix A are equal to the positive square roots of the non-zero eigenvalues of the $n \times n$ matrix A^*A (and AA^*). The eigenvectors of AA^* are the columns of U and the eigenvectors of A^*A are the columns of V.

0.5 Pseudoinverse

One of the most common computations needed in Machine Learning and particularly in *Regression*, is the computation of the (Moore-Penrose) pseudoinverse. It is fairly easy to obtain the pseudoinverse of a matrix once we have its Singular Value Decomposition:

$$A^+ = (UDV^*)^+ = VD^+U^* \tag{9}$$

where D^+ is the pseudo-inverse of D, also diagonal and with entries equal to the reciprocals of the non-zero singular values:

$$\sigma_i^+ = \begin{cases} \frac{1}{\sigma_i} & \text{if} \quad \sigma_i \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$