Single value decomposition and pseudo-inverses

Kristian Wichmann

December 12, 2016

1 Gramian matrices

Given a set of vectors $a_1, a_2, \dots, a_n \in \mathbb{R}^m$, the Gramian matrix is the traditionally matrix of inner products $\langle a_i, a_j \rangle$. If these vectors are collected into a $m \times n$ matrix A, this matrix can be expressed as $A^t A$. Here, we will use the term for any matrix in this form. By starting out with the transpose instead, this means that AA^t is also a Gramian, with dual results.

Theorem 1.1. *If $A \in \mathbb{R}^{m\times n}$, then $A^t A$ is symmetric and positive semi-definite. Iff A has rank m, $A^t A$ is positive definite.*

Proof. $(A^t A)^t = A^t (A^t)^t = A^t A$ shows symmetry. positive semi-definiteness, let $x \in \mathbb{R}^n$. Then:

$$x^t A^t A x = \langle Ax, Ax \rangle = ||Ax||^2 \tag{1.1}$$

As a norm, this is greater than or equal to zero. Hence $A^t A$ is positive semi-definite. If A has rank m the map $x \mapsto Ax$ has a trivial kernel by the rank-kernel theorem. Which means only the zero vector is mapped to zero, and hence $A^t A$ is positive definite. If the rank is less than m, the kernel is non-trivial and positive definiteness cannot be true. □

2 Single value decomposition

Let $A \in \mathbb{R}^{m\times n}$. Since $A^t A$ is symmetric, it is diagonalizable. So there is an orthogonal $n \times n$ matrix O such that $A^t A = ODO^t$, where D is a diagonal matrix of eigenvalues.

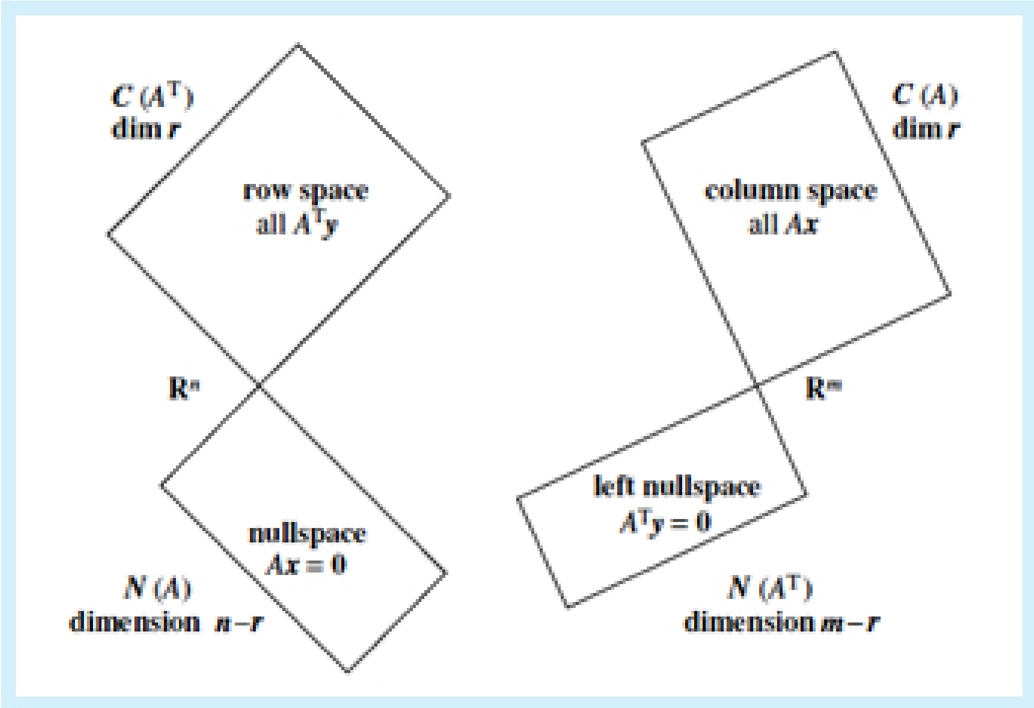

Figure 1: Visualization of dimensionality for the rank-nullity theorem

3 Generalized inverses

For an invertible matrix A, it's obviously true that:

$$AA^{-1}A = A \tag{3.1}$$

If A is not invertible, we may still define a *generalized inverse* A^g as a matrix that satisfies the same equation:

$$AA^gA = A \tag{3.2}$$

If A^g further satisfies:

$$A^gAA^g = A^g, \tag{3.3}$$

it is called a *reflexive generalized inverse*.

3.1 Left and right inverses

If $A \in \mathbb{R}^{m\times n}$ has rank n, then the null space is trivial, and hence the corresponding linear transformation is injective. This means that the equation $Ax = b$ may or may not have a solution, but if it exists, it's unique. The matrix A^tA has rank n as well, and hence is invertible. This can be used to construct a left inverse:

$$A_L^{-1} = (A^tA)^{-1}A^t, \qquad A_L^{-1}A = (A^tA)^{-1}A^tA = I_n \tag{3.4}$$

Similarly, if $A \in \mathbb{R}^{m\times n}$ has rank m, then the image space is all of $\mathbb{R}^m$, and hence the corresponding linear transformation is surjective. This means that the equation $Ax = b$ always has a solution, and it may have infinitely many. The matrix AA^t has rank m as well, and hence is invertible. Analogously, we can use this to construct a right inverse:

$$A_R^{-1} = A^t(AA^t)^{-1}, \qquad AA_R^{-1} = AA^t(AA^t)^{-1} = I_m \tag{3.5}$$

Both of of these inverses (when they exist) satisfies equation 3.2. They also satisfy 3.3. For instance:

$$A_L^{-1}AA_L^{-1} = (A^tA)^{-1}A^tA(A^tA)^{-1}A^t = (A^tA)^{-1}A^t = A_L^{-1} \tag{3.6}$$

4 The Moore-Penrose pseudoinverse

The *Moore-Penrose pseudoinverse* or simply the pseudoinverse of a real matrix A is the reflexive, generalized inverse A^+ which also satisfies:

$$(AA^+)^t = AA^+, \qquad (A^+A)^t = A^+A \tag{4.1}$$

In other words, for which AA^+ and A^+A are symmetrical.