DECOMPOSITIONS · SVD

What the SVD reveals

Every entry comes from the same three factors. What changes is which one is being read — the singular values for size, the columns of U and V for structure, and all three together for geometry.

7
READINGS

FROM THE SINGULAR VALUES 3

1 Rank

`strictly positive only`

$r = \text{count of } \sigma_i > 0$

The most numerically reliable rank there is. Row reduction decides rank by comparing entries against zero, which floating point makes arbitrary; the singular values instead show a gap, and where that gap falls is a judgement the numbers themselves support.

2 Norms

`read straight off the list`

$\|A\|_2 = \sigma_1, \quad \|A\|_F = \sqrt{\sum \sigma_i^2}$

The largest singular value is the most any unit vector is stretched. The Frobenius norm is the whole list in quadrature — so both common matrix norms are functions of the same numbers.

3 Condition number

A `of full rank`

$\kappa(A) = \sigma_1 / \sigma_r$

The ratio of most to least stretched. A large κ means the matrix is nearly singular in some direction, and small changes to $\mathbf{b}$ produce large changes to the solution — which no determinant reports, since a matrix can have $\det = 1$ and be badly conditioned.

FROM THE COLUMNS OF U AND V 2

4 The four subspaces

`orthonormal bases, all four at once`

partition U and V at index r

First r columns of V span the row space and the rest the null space; first r of U span the column space and the rest the left null space. The only method giving **all four** orthonormally from one computation.

5 Pseudoinverse

`reciprocate the nonzero σ, leave zeros alone`

$A^{+} = V \Sigma^{+} U^{\mathrm{T}}$

Defined for every matrix, square or not, invertible or not. When A is invertible it coincides with A^{-1}; otherwise it returns the least-squares solution of minimum norm, which is the sense in which it is the closest thing to an **inverse**.

FROM ALL THREE FACTORS 2

6 Low-rank approximation

`optimal for every unitarily invariant norm`

$A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}}$

Truncating the outer-product sum gives the best rank-k approximation there is — the Eckart–Young theorem, and it is why the SVD underlies image compression and principal component analysis. The error is exactly σ_{k+1}, so the singular values say in advance how much is lost.

7 Geometric action

every matrix, no hypotheses

$A = U\Sigma V^{\mathrm{T}}$ — rotate, stretch, rotate

Read right to left: V^{T} rotates, Σ scales along axes, U rotates again. Every linear map is those three steps, which is the claim that makes the factorization worth having as geometry rather than only as algebra.

No other factorization answers this many questions, and none exists for every matrix. That combination is why the SVD is the one to reach for when the matrix is rectangular, rank-deficient, or simply unknown — the conditions the **other decompositions** require are exactly the ones it does without.