Bringing the SVD to Life

By Gilbert Strang

The singular value decomposition is wonderful, but it is not so easy to teach. In desperation, I usually factor some low-rank matrices. Readers might be interested in two full-rank examples that connect the SVD to the derivative of sin kx. I believe they could help.

Symmetric positive definite matrices are one of the highlights of a linear algebra course. The equivalence of positive pivots, positive eigenvalues, and positive determinants brings the whole subject together. We need the matrix $A^{\mathrm{T}}A$ for least squares and projections, and it is the key to the SVD.

The singular value decomposition is the other half of this high point. $A = U\Sigma V^{\mathrm{T}}$ gives perfect bases for the four fundamental subspaces. The first r columns of U and V are orthonormal bases for the column space and row space of A. The remaining $m - r$ and $n - r$ columns are orthonormal bases for the nullspaces of A^{T} and A. The key equation $AV = U\Sigma$ says that A is diagonalized by these bases: Each $Av_k = \sigma_k u_k$. The positive eigenvalues of $A^{\mathrm{T}}A$ and AA^{T} are the squares of the singular values σ_k in the diagonal matrix Σ. Nothing better could be imagined, except that examples tend to be small and artificial and quite unimpressive to students.

That is the downside of the SVD. We can invent matrices and compute $U\Sigma V^{\mathrm{T}}$. We can identify Σ and V from the eigenvalues and eigenvectors in $A^{\mathrm{T}}Av_k = (\sigma_k)^2 v_k$. And we can produce $u_k = Av_k/\sigma_k$ as eigenvectors of AA^{T}. The quick proof is so typical of linear algebra: *Multiply by A and move the parentheses in* $(AA^{\mathrm{T}})(Av_k) = (\sigma_k)^2(Av_k)$. But bringing the factorization to life is much harder than its proof, especially if time is running out on the course.

My hope lies with a simple difference matrix. Put 1's on the diagonal and –1's on the subdiagonal of an $n + 1$ by n matrix A. Then $A^{\mathrm{T}}A$ is the –1,2,–1 second difference matrix `toeplitz([2 -1 zeros(1,n-2)])`. Its pivots, 2/1, 3/2, 4/3, . . . , are perfect examples for elimination, giving 2,3,4, . . . as the principal determinants. But the SVD requires eigenvectors. My point is that the eigenvectors of these matrices are attractive and important too.

Expressed in words, $Av_k = \sigma_k u_k$ says that *the difference of a discrete sine is a discrete cosine*. With $(n + 1)h = \pi$, the sine and cosine eigenvectors are

$$v_k = (\sin kh, \sin 2kh, \ldots, \sin nkh)$$
$$u_k = (\cos kh/2, \cos 3kh/2, \ldots, \cos (n + .5)kh)$$

A typical difference is $\sin 2kh - \sin kh$. This equals $2 \sin kh/2 \cos 3kh/2$. Identities like that give the scalar form of the SVD, and $Av_k = 2 \sin kh/2 u_k$ is the vector form. Far better is the matrix form, with the discrete cosine and sine transforms used as U and V:

SVD Example 1: $A = (\mathrm{DCT})\ \Sigma\ (\mathrm{DST})^{\mathrm{T}}$

The cosine vector $u_{n+1} = (1, \ldots, 1)$ from the nullspace of A^{T} is conventionally placed in column 0 (displeasing MATLAB) of the DCT matrix. All columns of the DCT and DST matrices are normalized to unit vectors (luckily, $\|u_k\| = \|v_k\|$ above).

That matrix $A^{\mathrm{T}}A$ builds in the Dirichlet con-ditions $v(0) = v(\pi) = 0$. The matrix AA^{T} with $(1, \ldots, 1)$ in its nullspace changes to Neumann conditions $u'(0) = u'(\pi) = 0$.

By dropping the last row of A to form a square matrix B, we get a second example that combines Dirichlet at one end with Neumann at the other. The 1,1 entry of $B^{\mathrm{T}}B$ and the n,n entry of BB^{T} change from 2 to 1. The effective meshlength H comes from $(n + .5)H = \pi$. Each vector u_k just reverses the components of v_k:

$$v_k = (\sin (k - .5)H, \sin 2(k - .5)H, \ldots)$$

The same identities produce $Bv_k = [2 \sin (k - .5)H/2]\ u_k$ with the new v and u. The matrix form has the "half-point" discrete cosine and sine transforms:

SVD Example 2: $B = (\mathrm{DCTH})\ \Sigma\ (\mathrm{DSTH})$

The real magic is that the discrete eigenvectors just sample the sine and cosine eigenfunctions of the continuous problems. Once again, and probably forever, Fourier sets the best example.

Gilbert Strang is a professor of mathematics at MIT.