2019-11-01

1 Symmetric eigenvalue basics

The *symmetric (Hermitian) eigenvalue problem* is to find nontrivial solutions to

$$Ax = x\lambda$$

where $A = A^*$ is symmetric (Hermitian). The symmetric eigenvalue problem satisfies several properties that we do not have in the general case:

- All eigenvalues are real.
- There are no non-trivial Jordan blocks.
- Eigenvectors associated with distinct eigenvalues are orthogonal.

It is worthwhile to make some arguments for these facts, drawing on ideas we have developed already:

- For any v, $v^*Av = v^*A^*v = \bar{v^*Av}$, so v^*Av must be real; and we can write any eigenvalue as v^*Av where v is the corresponding eigenvector (normalized to unit length).
- If $(A - \lambda I)^2 v = 0$ for $\lambda \in \mathbb{R}$ and $v \neq 0$, then
$$0 = v^*(A - \lambda I)^2 v = \|(A - \lambda I)v\|^2 = 0;$$
and so $(A - \lambda I)v = 0$ as well. But if λ is associated with a Jordan block, there must be $v \neq 0$ such that $(A-\lambda I)^2 v = 0$ and $(A-\lambda I)v \neq 0$.
- If $\lambda \neq \mu$ are eigenvalues associated with eigenvectors u and v, then
$$\lambda u^* v = u^* A v = \mu u^* v.$$
But if $\lambda \neq \mu$, then $(\lambda - \mu)u^*v = 0$ implies that $u^*v = 0$.

We write the complete eigendecomposition of A as

$$A = U\Lambda U^*$$

where U is orthogonal or unitary and Λ is a real diagonal matrix. This is simultaneously a Schur form and a Jordan form.

More generally, if $\langle \cdot, \cdot \rangle$ is an inner product on a vector space, the *adjoint* of an operator A on that vector space is the operator A^* s.t. for any v, w

$$\langle Av, w \rangle = \langle v, A^* w \rangle.$$

If $A = A^*$, then A is said to be *self-adjoint*. If a matrix A is self-adjoint with respect to the M-inner product $\langle v, w \rangle_M = w^* M v$ where M is Hermitian positive definite, then $H = MA$ is also Hermitian. In this case, we can rewrite the eigenvalue problem

$$Ax = x\lambda$$

as

$$Hx = MAx = Mx\lambda.$$

This gives a *generalized* symmetric eigenvalue problem[1]. A standard example involves the analysis of reversible Markov chains, for which the transition matrix is self-adjoint with respect to the inner product defined by the invariant measure.

For the generalized problem involving the matrix pencil (H, M), all eigenvalues are again real and there is a complete basis of eigenvectors; but these eigenvectors are now M-orthogonal. That is, there exists U such that

$$U^* H U = \Lambda, \quad U^* M U = I.$$

Generalized eigenvalue problems arise frequently in problems from mechanics. Note that if $M = R^T R$ is a Cholesky factorization, then the generalized eigenvalue problem for (H, M) is related to a standard symmetric eigenvalue problem

$$\hat{H} = R^{-T} H R^{-1};$$

if $\hat{H}x = x\lambda$, then $Hy = My\lambda$ where $Ry = x$. We may also note that $R^{-1}\hat{H}R = M^{-1}H$; that is $\hat{H}$ is related to $A = M^{-1}H$ by a similarity transform. Particularly for the case when M is large and sparse, though, it may be preferable to work with the generalized problem directly rather than converting to a standard eigenvalue problem, whether or not the latter is symmetric.

[1]The case where M is allowed to be indefinite is not much nicer than the general nonsymmetric case.

The singular value decomposition may be associated with several different symmetric eigenvalue problems. Suppose $A \in \mathbb{R}^{n \times n}$ has the SVD $A = U \Sigma V^T$; then

$$
\begin{aligned}
A^T A &= V \Sigma^2 V^T \\
A A^T &= U \Sigma^2 U^T \\
\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} U & U \\ V & -V \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix} \begin{bmatrix} U & U \\ V & -V \end{bmatrix}^T.
\end{aligned}
$$

The picture is marginally more complicated when A is rectangular — but only marginally.

2 Variational approaches

The Rayleigh quotient plays a central role in the theory of the symmetric eigenvalue problem. Recall that the Rayleigh quotient is

$$\rho_A(v) = \frac{v^* A v}{v^* v}.$$

Substituting in $A = U \Lambda U^*$ and (without loss of generality) assuming $w = U^* v$ is unit length, we have

$$\rho_A(v) = \sum_{i=1}^N \lambda_i |w_i|^2, \quad \text{with } \sum_{i=1}^N |w_i|^2 = 1.$$

That is, the Rayleigh quotient is a weighted average of the eigenvalues. Maximizing or minimizing the Rayleigh quotient therefore yields the largest and the smallest eigenvalues, respectively; more generally, for a fixed A,

$$\delta \rho_A(v) = \frac{2}{\|v\|^2} \, \delta_v^* \left(Av - v \rho_A(v) \right),$$

and so at a stationary v (where all derivatives are zero), we satisfy the eigenvalue equation

$$Av = v \rho(A).$$

The eigenvalues are the stationary values of ρ_A; the eigenvectors are stationary vectors.

The Rayleigh quotient is homogeneous of degree zero; that is, it is invariant under scaling of the argument, so $\rho_A(v) = \rho_A(\tau v)$ for any $\tau \neq 0$. Hence, rather than consider the problem of finding stationary points of ρ_A generally, we might restrict our attention to the unit sphere. That is, consider the Lagrangian function

$$L(v, \lambda) = v^* A v - \lambda(v^* v - 1);$$

taking variations gives

$$\delta L = 2\delta v^*(Av - \lambda v) - \delta\lambda(v^* v - 1)$$

which is zero only if $Av = \lambda v$ and v is normalized to unit length. In this formulation, the eigenvalue is identical to the Lagrange multiplier that enforces the constraint.

The notion of a Rayleigh quotient generalizes to pencils. If M is Hermitian and positive definite, then

$$\rho_{A,M}(v) = \frac{v^* A v}{v^* M v}$$

is a weighted average of generalized eigenvalues, and the stationary vectors satisfy the generalized eigenvalue problem

$$Av = Mv\rho_{A,M}(v).$$

We can also restrict to the ellipsoid $\|v\|_M^2 = 1$, i.e. consider the stationary points of the Lagrangian

$$L(v, \lambda) = v^* A v - \lambda(v^* M v - 1),$$

which again yields a generalized eigenvalue problem.

The analogous construction for the SVD is

$$\phi(u, v) = \frac{u^* A v}{\|u\|\|v\|}$$

or, thinking in terms of a constrained optimization problem,

$$L(u, v, \lambda, \mu) = u^* A v - \lambda(u^* u - 1) - \mu(v^* v - 1).$$

Taking variations gives

$$\delta L = \delta u^* (Av - 2\lambda u) + \delta v^* (A^* u - 2\mu v) - \delta\lambda (u^* u - 1) - \delta\mu (v^* v - 1),$$

and so $Av \propto u$ and $A^* u \propto v$. Combining these observations gives $A^* Av \propto v$, $AA^* u \propto u$, which we recognize as one of the standard eigenvalue problem formulations for the SVD, with the squared singular values as the constant of proportionality.