Eigensystems for 3×3 Symmetric Matrices

David Eberly
Magic Software, Inc.
http://www.magic-software.com

Created: October 9, 2002
Modified: October 10, 2002 (noted that matrix U is adjoint of M)
Modified: November 18, 2002 (case $Q = 0$ missing the $c_2/3$ terms)
Modified: November 19, 2002 (another typo, $Q < 0$ case had missing power of 1/3 on ρ)
Modified: November 20, 2002 (wrote the wrong roots again for $Q = 0$, how hard is it to fix a couple of typos)

A standard result from linear algebra is that an $n \times n$ symmetric matrix A with real–valued entries must have n real–valued and unit–length eigenvectors $\vec{v}_1$ through $\vec{v}_n$ that are mutually orthogonal. Each vector satisfies $A\vec{v}_i = \lambda_i \vec{v}_i$ where λ_i is the eigenvalue associated with the eigenvector. The eigenvalues are not necessarily distinct. If the eigenvectors are stored as the columns of a matrix $P = [\vec{v}_1 \ \cdots \ \vec{v}_n]$, this matrix is orthogonal. Also store the eigenvalues as the diagonal entries of a diagonal matrix $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. The n eigenvector equations can be written as a single matrix equation $AP = PD$. Since P is orthogonal, $P^{-1} = P^{\mathrm{T}}$, so equivalently $A = PDP^{\mathrm{T}}$. Various iterative numerical methods may be applied to factoring A. One method uses Jacobi transformations to approximate P by a composition of rotation matrices $Q = Q_1 \cdots Q_k$ with k large enough so that $Q^{\mathrm{T}}AQ$ is effectively diagonal. Another method uses either Givens reductions or Householder reductions to obtain a matrix Q in a fixed number of steps so that $B = Q^{\mathrm{T}}AQ$ is tridiagonal. The matrix B is then factored to $R^{\mathrm{T}}DR$ using an iterative scheme such as the QR or QL algorithms. The implementations of these methods are designed to be accurate and robust.

For 3×3 symmetric matrices, it is possible to avoid the iterative methods. Let $A = [a_{ij}]$ for $1 \le i \le 3$ and $1 \le j \le 3$ with $a_{ji} = a_{ij}$. Generally, an eigenvalue λ of the matrix is a root to the polynomial equation $\det(A - \lambda I) = 0$ where I is the identity matrix and det denotes the determinant function. In the 3×3 case, the polynomial equation is

$$0 = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{12} & a_{22} - \lambda & a_{23} \\ a_{13} & a_{23} & a_{33} - \lambda \end{bmatrix} = -\lambda^3 + c_2\lambda^2 - c_1\lambda + c_0$$

where

$$\begin{aligned} c_0 &= a_{11}a_{22}a_{33} + 2a_{12}a_{13}a_{23} - a_{11}a_{23}^2 - a_{22}a_{13}^2 - a_{33}a_{12}^2 \\ c_1 &= a_{11}a_{22} - a_{12}^2 + a_{11}a_{33} - a_{13}^2 + a_{22}a_{33} - a_{23}^2 \\ c_2 &= a_{11} + a_{22} + a_{33} \end{aligned}$$

As a side note, c_0 is the determinant of A and c_2 is the trace of A. The polynomial is cubic, so the standard closed–form formulas may be used to compute the roots. Set $a = (3c_1 - c_2^2)/3$, $b = (-2c_2^3 + 9c_1c_2 - 27c_0)/27$, and $Q = b^2/4 + a^3/27$. If $Q > 0$, there is one root. For this to occur it must be that $A = \lambda I$, so there is no need to use a root solver here. Keep a "hint" in your program whether or not A is a multiple of the identity matrix. If $Q = 0$, there are two distinct roots,

$$\begin{aligned}\lambda_1 &= c_2/3 + (b/2)^{1/3}\\ \lambda_2 &= c_2/3 + (b/2)^{1/3}\\ \lambda_3 &= c_2/3 - 2(b/2)^{1/3}\end{aligned}$$

If $Q < 0$, there are three distinct roots. Let $\theta = \operatorname{atan2}(\sqrt{-Q}, -b/2)$ and $\rho = \sqrt{(b/2)^2 - Q}$. The roots are

$$\begin{aligned}\lambda_1 &= c_2/3 + 2\rho^{1/3}\cos(\theta/3)\\ \lambda_2 &= c_2/3 - \rho^{1/3}(\cos(\theta/3) + \sqrt{3}\sin(\theta/3))\\ \lambda_3 &= c_2/3 - \rho^{1/3}(\cos(\theta/3) - \sqrt{3}\sin(\theta/3))\end{aligned}$$

Define $M = A - \lambda I = [m_{ij}]$. For a specific eigenvalue λ, a corresponding eigenvector $\vec{v} \neq \vec{0}$ is obtained by solving $M\vec{v} = \vec{0}$. There must be a nonzero solution since $\det(M) = 0$. The linear system may be solved by Gaussian elimination. This can be done symbolically, but the outcome depends on the multiplicity of λ as a root of the polynomial. Also, let $\vec{m}_i$ denote row i of M. Observe that the equation $M\vec{v} = \vec{0}$ says that $\vec{v}$ is perpendicular to each row of M, $\vec{m}_i \cdot \vec{v} = 0$. The constructions presented here generate unit–length eigenvectors. If you do not need unit–length, you can skip the normalization steps.

Case 1: λ has multiplicity 1.

The rank of M must be 2, so two rows of the matrix are linearly independent. Since $\vec{v}$ is perpendicular to both rows, $\vec{v}$ must be in the direction of the cross product of those rows. However, you do not know in advance which two rows are linearly independent. Compute all the cross products $\vec{u}_1 = \vec{m}_2 \times \vec{m}_3$, $\vec{u}_2 = \vec{m}_3 \times \vec{m}_1$, and $\vec{u}_3 = \vec{m}_1 \times \vec{m}_2$. Stored as the columns of a matrix $U = [\vec{u}_1\ \vec{u}_2\ \vec{u}_3] = [u_{ij}]$, we have

$$U = \begin{bmatrix} m_{22}m_{33} - m_{23}^2 & m_{13}m_{23} - m_{12}m_{33} & m_{12}m_{23} - m_{13}m_{22} \\ m_{13}m_{23} - m_{12}m_{33} & m_{11}m_{33} - m_{13}^2 & m_{12}m_{13} - m_{23}m_{11} \\ m_{12}m_{23} - m_{13}m_{22} & m_{13}m_{12} - m_{23}m_{11} & m_{11}m_{22} - m_{12}^2 \end{bmatrix}$$

Observe that U is symmetric, so only the six entries u_{11}, u_{12}, u_{13}, u_{22}, u_{23}, and u_{33} need to be computed. Keep track of the maximum of $|u_{ij}|$ and the column c in which it occurs. The unit–length eigenvector is selected as $\vec{v} = \vec{u}_c/|\vec{u}_c|$. Also note that $U = \operatorname{adj}(M)$, the adjoint of M and is the transpose of the matrix of cofactors. For any square matrix M it is true that $M\operatorname{adj}(M) = \operatorname{adj}(M)M = \det(M)I$. If $\det(M) \neq 0$, M is invertible and the inverse is $M^{-1} = \operatorname{adj}(M)/\det(M)$. In our current case, M is not invertible since $\det(M) = 0$. Consequently, $M\operatorname{adj}(M) = 0$ (the zero matrix).

Case 1: λ has multiplicity 2.

The rank of M must be 1, so only one row of the matrix is linearly independent, call it $\vec{m}_r$ There must be two orthogonal and unit–length eigenvectors that solve $(A - \lambda I)\vec{v} = \vec{0}$. These are chosen to be in the plane whose normal is $\vec{m}_r$. When computing the six distinct entries of M, keep track of the maximum of $|m_{ij}|$ $(i \leq j)$ and the row r and column c in which it occurs. Select vectors $\vec{w}_1$ and $\vec{w}_2$ according to the table below. Both are eigenvectors, so the unit–length eigenvectors are $\vec{v}_1 = \vec{w}_1/|\vec{w}_1|$ and $\vec{v}_2 = \vec{w}_2/|\vec{w}_2|$:

r	c	$\vec{w}_1$	$\vec{w}_2$
1	$1,2$	$(-m_{12}, m_{11}, 0)$	$(-m_{13}m_{11}, -m_{13}m_{12}, m_{11}^2 + m_{12}^2)$
1	3	$(m_{13}, 0, -m_{11})$	$(-m_{12}m_{11}, m_{11}^2 + m_{13}^2, -m_{12}m_{13})$
2	$2,3$	$(0, -m_{23}, m_{22})$	$(m_{22}^2 + m_{23}^2, -m_{12}m_{22}, -m_{12}m_{23})$
3	3	$(0, -m_{33}, m_{23})$	$(m_{23}^2 + m_{33}^2, -m_{13}m_{23}, -m_{13}m_{33})$

Case 3: λ has multiplicity 3.

It must be the that $A = \lambda I$, so eigenvectors are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.