Mathematical Physics II

PHYS 502

Solution to Problem Set # 1

1. Here is a 3×2 matrix:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

Diagonalize it.

Solution: The Singular Value Decomposition (SVD) was discovered about a dozen different times in about a dozen different areas of research. It says

$$T_{ir} = \sum_{\alpha=1,min(p,q)} u(i,\alpha)\lambda_\alpha v(r,\alpha) \qquad 1 \leq i \leq p, \quad 1 \leq r \leq q$$

Here $u(i,\alpha)$ is the ith component of the α eigenvector of the $p \times p$ real symmetric matrix $M3_{ij} = TT^t = T_{ir}T_{jr}$ whose nonzero eigenvalues are λ_α^2 and $v(r,\alpha)$ is the rth component of the α eigenvector of the real symmetric $q \times q$ matrix $M2_{rs} = T^tT = T_{ir}T_{is}$ whose nonzero eigenvalues are the same λ_α^2:

$$M3 = TT^t = \begin{bmatrix} 5 & 11 & 17 \\ 11 & 25 & 39 \\ 17 & 39 & 61 \end{bmatrix} \qquad M2 = T^tT = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}$$

The eigenvalues of $M3$ are $\lambda_\alpha^2 = 90.735..., 0.264..., 0$ and of $M2$ are $\lambda_\alpha^2 = 90.735..., 0.264...$. It is not a clever idea to diagonalize both matrices. There are two reasons. (1) Often one of p, q is much larger than the other. (2) There are phase relations among the eigenvectors, and separate diagonalization destroys these phase relations. It is therefore useful to compute the eigenvectors $u(i,\alpha)$ by taking the inner product of T with the eigenvectors of the smaller matrix:

$$T_{i,r}v(r,\beta=1) = \left\{ \sum_\alpha u(i,\alpha)\lambda_\alpha v(r,\alpha) \right\} v(r,1) = \lambda_1 u(i,1)$$

$$T_{i,r} v(r, \beta = 2) = \left\{ \sum_{\alpha} u(i,\alpha) \lambda_\alpha v(r,\alpha) \right\} v(r,2) = \lambda_2 u(i,2)$$

For the example at hand

$$\lambda_1 u(i,1) = \begin{bmatrix} 2.187 \\ 4.993 \\ 7.799 \end{bmatrix} \qquad \lambda_2 u(i,2) = \begin{bmatrix} -0.454 \\ -0.124 \\ 0.206 \end{bmatrix}$$

with $\lambda_1 = \sqrt{\lambda_1^2} = 9.525$ and $\lambda_2 = \sqrt{\lambda_2^2} = 0.514$.

Creeping up on a solution we construct $u(i,1)\lambda_1 v(r,1)$ and $u(i,2)\lambda_2 v(r,2)$:

$$u(i,1)\lambda_1 v(r,1) = \begin{bmatrix} 1.353 & 1.714 \\ 3.090 & 3.914 \\ 4.827 & 6.114 \end{bmatrix} \qquad u(i,2)\lambda_2 v(r,2) = \begin{bmatrix} -0.355 & 0.281 \\ -0.097 & 0.076 \\ 0.161 & -0.127 \end{bmatrix}$$

Note that the sum of these two matrices is T up to roundoff error, and the difference gets smaller as more and more decimal digits are retained, and goes to zero eventually.

Note also that only the first matrix $u(i,1)\lambda_1 v(r,1)$ is already a good approximation to T. This is because the first eigenvalue is so much larger than the second. Quantitatively, this first matrix contains $\frac{90.735}{90.735+0.264} = 0.997 = 99.7\%$ of the information contained in T, estimated in a least sqaures sense.