Lecture 8 — February 8

Lecturer: Venkat Anantharam *Scribe: none*

This lecture was not scribed. Most of the calculations done during this lecture are available in the textbook.

We started out by discussing the structure of symmetric matrices. An $n \times n$ real matrix A is called *symmetric* if it equals its transpose, i.e. $A = A^T$.

Every eigenvalue of a symmetric matrix is real. This may be seen by considering it as a complex matrix. Let λ be an eigenvalue (possibly complex) and v an associated eigenvector (possibly complex), so we have $Av = \lambda v$. Consider the expression $v^* A v$, where v^* denotes the complex conjugate transpose of v. Since $v^* A^* = v^* A$ (because A^*, the complex conjugate transpose of A, equals A^T by virtue of A being real, and also equals A, by virtue of A being symmetric), we have $v^* A = \lambda^* v^*$ by taking the complex conjugate transpose of the equation $Av = \lambda v$, where λ^* denotes the complex conjugate of λ. Now we can write

$$\lambda^* v^* v = v^* A v = \lambda v^* v \ ,$$

from which it follows that $\lambda^* = \lambda$, i.e. λ is real (because v is not the zero vector).

If v_1 and v_2 are eigenvector of the symmetric matrix A corresponding to distinct eigenvalues λ_1 and λ_2 respectively, then they are orthogonal. To see this, using the already proved fact that both λ_1 and λ_2 are real, we may write

$$\lambda_2 v_2^T v_1 = v_2 A v_1 = \lambda_1 v_2^T v_1 \ ,$$

where in one case we used $v_2^T A^T = v_2^T A = \lambda_2 v_2^T$ and in the other case we used $Av_1 = \lambda_1 v_1$. From this it follows that $v_2^T v_1 = 0$, as claimed, because $\lambda_1 \neq \lambda_2$.

Every symmetric matrix has a complete basis of eigenvectors. This is a consequence of the fact that $\mathcal{N}((\lambda I - A)^2) = \mathcal{N}(\lambda I - A)$ for any eigenvalue λ of A. To see this, suppose $v \in \mathcal{N}((\lambda I - A)^2)$ is nonzero and consider

$$v^T (\lambda I - A)^2 v = 0 \ .$$

Since $(\lambda I - A)^T = (\lambda I - A)$, this can be written as

$$((\lambda I - A)v)^T (\lambda I - A) v = 0 \ ,$$

from which it follows that $v \in \mathcal{N}(\lambda I - A)$.

From this discussion it follows that there is an orthogonal $n \times n$ matrix T (each of its columns will be an eigenvector of A, each column will have squared norm 1, and any distinct pair of its columns is orthogonal) such at $T^{-1}AT$ is a diagonal matrix with its diagonal entries being the eigenvalues of A, each with multiplicity equal to its multiplicity as a root of the characteristic polynomial of A. Note that $T^{-1} = T^T$ (this is what it means for T to be orthogonal).

A symmetric matrix is called *positive semidefinite* or *nonnegative definite* if all its eigenvalues are nonnegative. It is called *positive definite* if all its eigenvalues are strictly positive.

We discussed the fully observed finite horizon problem of controlling the linear system

$$x_{k+1} = A_k x_k + B_k u_k + w_k \ , \quad k = 0, 1, \ldots, N-1$$

with the quadratic cost criterion of minimizing (informally)

$$E[\sum_{k=0}^{N-1} (X_k^T Q_k X_k + U_k^T R_k U_k) + X_N^T Q_N X_N] \ ,$$

where each Q_k, $k = 0, 1, \ldots, N$ is a positive semidefinite symmetric matrix, and each R_k, $k = 0, 1, \ldots, N-1$ is a positive definite symmetric matrix. Here $x_k \in \mathbf{R}^n$ and $u_k \in \mathbf{R}^m$ for each k. w_k is a zero mean vector noise with finite autocovariance matrix. See Section 4.1 of the text.

We proved that the optimal control strategy can be taken as a linear feedback strategy based on the current state, with the matrices defining the policy coming from the Riccati equation, as in eqns. (4.3) and (4.4) and the equation that precedes them, in the textbook.

The problem formulation is motivated by trying to make the state track zero, with deviations from the desired trajectory being subject to a direction dependent quadratic penalty (given by the matrices Q_k) and control effort also being penalized by a direction dependent quadratic penalty (given by the matrices R_k). Further, the natural assumption is made that there are no *freebie* control directions (this is what it means to assume that each R_k is positive definite rather than just positive semidefinite).