Lecture 4-26

We begin by going over the midterm. The first problem is a straightforward application of Lagrange multipliers; equating the gradient of volume to a multiple of the gradient of surface area and taking the variables r, h in that order, we get $(2\pi rh, \pi r^2) = \lambda(2\pi h + 4\pi r, 2\pi r)$, from which we get $2h/r = 2 + (h/r), (h/r) = 2$, at the unique critical point. This must correspond to maximum volume for fixed surface area, since the volume can be arbitrarily small for a given area. Hence *the height should be twice the radius of the base*. In the next problem, we observe that the given vectors are nonzero multiples of $(1, 1, 0), (1, 0, 1)$, and $(0, 1, 1)$. By taking the sum of these last vectors, we see that $(1, 1, 1)$ lies in their span; subtracting each of these last vectors in turn from $(1, 1, 1)$, we see that the vectors span $\mathbb{R}^3$. Since the dimension of $\mathbb{R}^3$ is 3, *the given vectors indeed form a basis*. In the third problem, we find that we must have $yz = xz = xy = -1$ at a critical point; but then taking the product of these equations yields $x^2y^2z^2 = -1$, which is impossible. Hence *there are no critical points and thus no local maxima or minima for this function*. In the fourth problem, applying row and column operations to clear out the first row and column, we get $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 8/3 & 2/3 \\ 0 & 2/3 & 8/3 \end{pmatrix}$; doing the same with the second row and column we get $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 8/3 & 0 \\ 0 & 0 & 5/2 \end{pmatrix}$. Since all diagonal entries are positive, *the given matrix is positive definite*. Finally, for the last problem, on bringing the given matrix to echelon form, we get $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & a-1 \end{pmatrix}$, whence we get three pivots if and only if the system always has a solution, or if and only if $a \neq 1$.

We now officially define an *eigenvector* of a square matrix A to be a nonzero vector $\vec{v}$ such that the product $A\vec{v} = \lambda\vec{v}$ for some $\lambda \in \mathbb{R}$ when $\vec{v}$ is written as a column vector; we call λ the corresponding *eigenvalue* and sometimes describe it more precisely as a λ-eigenvector. Similarly, we say that λ is an eigenvector of a linear transformation $f : V \to V$ if $f(v) = \lambda v$ for some $v \in V, v \neq 0$. The set of eigenvalues of a matrix or a transformation is called its *spectrum* (and behaves analogously to a chemical spectrum in certain ways). If λ is an eigenvalue of A, then the set of all λ-eigenvectors of A together with $\vec{0}$ is easily seen to be a vector space, called the λ-*eigenspace* of A; we use the same terminology for linear transformations. It is clear how to compute this space for a given A if λ is also given, as it will be the nullspace of the difference $A - \lambda I$. The hard part is finding a λ for which this space is nonzero, since we already know that most square matrices are invertible (so that their nullspaces are 0). More precisely, we now know that $A - \lambda I$ has a nonzero nullspace if iand only if it is singular, or if and only if $\det(A - \lambda I) = 0$. Thanks to the definition of the determinant, this last condition amounts to a polynomial condition on λ. More precisely, if A is $n \times n$, then there is a polynomial $p(x)$ of degree n and leading coefficient $(-1)^n$ such that $A - \lambda I$ is singular if and only if $p(\lambda) = 0$ (the coefficient $(-1)^n$ of the top-degree term λ^n in this polynomial arises from the product of the diagonal entries of $A - \lambda I$, which is one of the terms in $\det(A - \lambda I)$. We call this polynomial the *characteristic polynomial* of A. Now a polynomial of degree n over $\mathbb{R}$ or $\mathbb{C}$ always has exactly n roots, but there are two important caveats here: these roots can be complex even if the polynomial has real coefficients, and the roots have to be counted with multiplicity (so that the number of distinct roots might be strictly less than n). Hence *an $n \times n$ matrix always has at most n distinct eigenvalues, which can be complex even if the matrix has real entries.* As a simple example, take $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$. Here the characteristic polynomial $p(\lambda) = (1-\lambda)(4-\lambda) - (-2) = \lambda^2 - 5\lambda + 6 = (\lambda - 3)(\lambda - 2)$. Thus the eigenvalues of A are 3 and 2; to find corresponding eigenvectors, we find the nullspaces of $A - 3I$ and $A - 2I$. These spaces are spanned by $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, respectively; note that these vectors form a basis of the ambient vector space $\mathbb{R}^2$. (This is no coincidence, as we will see later that eigenvectors of a matrix corresponding to distinct eigenvalues are always linearly independent, so that in particular if an $n \times n$ matrix A has n distinct eigenvalues, then the corresponding eigenvectors always form a basis for the ambient vector space.)

Now the time has come to give a more detailed account of the how the matrix of a linear transformation $f : V \to V$ changes if the basis B of V used to produce this matrix is replaced by a different one B'. Fortunately, this is quite straightforward. Write the vectors in B' as linear combinations of the vectors in B and assemble a matrix P whose columns are the coefficients arising in this way (with the i-th column corresponding to the i-th vector in B'). Then it is quite easy to see that the matrix P is always invertible and in fact the columns of the inverse P^{-1} are the coefficients appearing when the vectors in B are written as combinations of the vectors in B'; moreover, by tracing through the definitions of P and P^{-1} we see that if A is the matrix of f with respect to B, then $P^{-1}AP$ is the matrix of f with respect to B'. As mentioned before, we therefore say that two square matrices M, N of the same size are *similar* if there is an invertible P with $N = P^{-1}MP$. One easily checks that M is similar to N if and only if N is similar to M, and if M is similar to N and N is similar to Q, then M is similar to Q (so that similarity is what is called an equivalence relation among matrices). Now if A admits a basis B of eigenvectors (corresponding to possibly different eigenvalues) then the matrix P whose columns are the vectors in B is easily seen to satisfy $P^{-1}AP = D$, a diagonal matrix; in fact the i-th diagonal entry of D is just the eigenvalue λ_i of the i-th vector in B. We say that A is *diagonalizable* in this situation. Many computations with diagonalizable matrices can be reduced to the corresponding computations with diagonal matrices, which are typically quite straightforward to carry out. For example, the k-th power of a diagonal matrix D with diagonal entries $d_1, \ldots, d_n$ is clearly another diagonal matrix with diagonal entries $d_1^k, \ldots, d_n^k$, for any integer k. But now when we compute $(P^{-1}DP)^2$ for some invertible matrix P, for example, we get $(P^{-1}DP)(P^{-1}DP) = P^{-1}D^2P$; more generally, we have $(P^{-1}DP)^k = P^{-1}D^kP$ for any integer k. Thus even large powers of diagonalizable matrices are quite simple to compute. In real-world applications, it frequently happens that a system evolves over time discretely, so that we can speak of its state on the k-th day (or some other time period) and this might depend only on its state on the $(k-1)$-st day. If the state on each day is encoded by a vector and the state on each day equal some fixed matrix times the vector encoding its state on the previous day, then by computing large powers of the matrix times the initial vector, we can predict the behavior of the system arbitrarily far in advance.