14.102, Math for Economists
Fall 2005
Lecture Notes, 9/29/2005

These notes are primarily based on those written by George Marios Angeletos for the Harvard Math Camp in 1999 and 2000, and updated by Stavros Panageas for the MIT Math for Economists Course in 2002. I have made only minor changes to the order of presentation, and added some material from Guido Kuersteiner's notes on linear algebra for 14.381 in 2002. The usual disclaimer applies; questions and comments are welcome.

Nathan Barczi
nab@mit.edu

7 Eigensystems: Applications

7.1 Linear Transformations

Given an $n \times m$ matrix A, consider the mapping $T : \mathbb{R}^m \to \mathbb{R}^n$ defined by $T(x) = Ax$ for all $x \in \mathbb{R}^m$. It is easy to check that this mapping is linear in the following sense:

Definition 172 *A mapping A of a vector space X into a vector space Y is called a* ***linear transformation*** *if*

$$A(x_1 + x_2) = Ax_1 + Ax_2 \quad \text{and} \quad A(cx) = cAx$$

for all $x, x_1, x_2 \in X$ and all scalars c.

Obviously, any matrix A induces a linear transformation. A fundamental result establishes a kind of converse, that any linear transformation can be uniquely represented by a matrix. Thus, we may think of matrices and linear transformations interchangeably. More precisely:

Theorem 173 *Given two vector spaces $\mathbb{X}$ and $\mathbb{Y}$ and any fixed bases $E = (e_1, ..., e_n)$ for $\mathbb{X}$ and $F = (f_1, ..., f_m)$ for $\mathbb{Y}$, there is a one-to-one correspondence between any linear transformation $T : \mathbb{X} \to \mathbb{Y}$ and a matrix $A = A(E, F)$. For given T, $n \times n$ E, $m \times m\, F$, the corresponding $A = [a_{ij}]$ is $m \times n$ and is given by projecting the image of E under T on F:*

$$A = F^{-1}T(E)$$

We may then let

$$T(x) = Ax \; \forall x \in \mathbb{X}$$

In more detail: For any $j = 1, ..., n$, take e_j from $\mathbb{X}$ and form its image under T. $T(e_j)$ is a vector in $\mathbb{Y}$. Since F is a basis for $\mathbb{Y}$, it must be the case that $T(e_j)$ can be written as a linear combination of the f_j's; that is,

$$\begin{aligned} T(e_j) &= a_{1j} f_1 + ... + a_{mj} f_m = F a_j \;\; \Rightarrow \\ a_j &= F^{-1} T(e_j) \end{aligned}$$

This means that a_j, the j-th column of A, consists simply of the coordinates of $T(e_j)$ on the basis F. More compactly,

$$\begin{aligned} T(E) &= FA \;\; \Rightarrow \\ A &= F^{-1} T(E) \end{aligned}$$

With A constructed so, take any $x \in \mathbb{X}$. Let the n-vector c be the coordinates of x on basis E; that is, $c = [c_j]$ is such that

$$x = Ec = c_1 e_1 + ... + c_n e_n$$

But then

$$T(x) = T(Ec) = T(c_1 e_1 + ... + c_n e_n)$$

and since T is a linear transformation

$$T(x) = c_1 T(e_1) + ... + c_n T(e_n) = T(E)c$$

Of course, $T(x) = T(E)c$ is a vector in $\mathbb{Y}$. Now using $T(E) = FA$, we get

$$T(x) = FAc \;\; \forall x \in \mathbb{X}$$

This means that Ac are the coordinates of $T(x)$ on the basis F. Thus, x is mapped to $T(x)$, but, for given E, F, x is equivalent to some c, and then $T(x)$ is equivalent to Ac. Thus $x \mapsto T(x)$ is equivalent to $c \mapsto Ac$. In this sense T is equivalent to A, and we may simply write $T(x) = Ax$, for x meant in E-coordinates and $T(x)$ in F-coordinates.

Exercise 174 *Show that the one-to-one correspondence between linear transformations T and matrices A preserves addition, scalar multiplication, and the zero element.*

Further, there is an immediate correspondence between inversion of linear transformations and matrix inversion:

Definition 175 *Given $\mathbb{X}$ and $\mathbb{Y}$, the mapping $T : \mathbb{X} \to \mathbb{Y}$ is invertible iff there is a mapping $T^{-1} : \mathbb{Y} \to \mathbb{X}$ such that, for all $x \in \mathbb{X}$,*

$$T^{-1}(T(x)) = x$$

Proposition 176 *Given $\mathbb{X}$ with basis E, $\mathbb{Y}$ with basis F, and $T : \mathbb{X} \to \mathbb{Y}$, let $A = F^{-1}T(E)$ be the equivalent matrix for T. Then, T is invertible if and only if A is invertible. If so, the inverse $T^{-1} : \mathbb{Y} \to \mathbb{X}$ is given by*

$$T^{-1}(y) = A^{-1} y \;\; \forall y \in \mathbb{Y}$$

Remark: Observe that the last result implies that, for $T : \mathbb{X} \to \mathbb{Y}$ to be invertible, its equivalent matrix A has to be square. (Or otherwise what would A^{-1} be?) Thus, $\mathbb{X}$ and $\mathbb{Y}$ have to share the same dimension. Without loss of generality, we may let $\mathbb{X} = \mathbb{Y}$.

Indeed, the identity matrix corresponds to the identity function

Definition 177 *We define the identity linear transformation by $I : \mathbb{X} \to \mathbb{X}$ by $I(x) = x \, \forall x \in \mathbb{X}$.*

Then, letting $\circ$ denote function composition, it follows by definition that the linear transformation $T : \mathbb{X} \to \mathbb{X}$ is invertible if and only if there is $T^{-1} : \mathbb{X} \to \mathbb{X}$ such that

$$T^{-1} \circ T = T \circ T^{-1} = I$$

Finally, the canonical form of a linear trasformation is provided by the canonical form of the corresponding matrix. That is, the canonical from of $x \mapsto y = Ax$ is $\tilde{x} \mapsto \tilde{y} = \Lambda \tilde{x}$ for $\Lambda = V^{-1}AV, \tilde{y} = V^{-1}y, \tilde{x} = V^{-1}x$.

7.2 Powers, Rank, Determinant: Using the Canonical Form

Let A be a diagonalizable matrix, and let Λ be the diagonal matrix of its eigenvalues and V that of its eigenvectors. We can easily show the following result:

Lemma 178 *For any diagonalizable $A = V \Lambda V^{-1}$, its **determinant** is given by the product of all its eigenvalues,*

$$|A| = |\Lambda| = \lambda_1 \lambda_2 ... \lambda_n$$

*It follows that A is nonsingular and **invertible** if and only if all its eigenvalues are nonzero,*

$$|A| \neq 0 \;\Leftrightarrow\; \lambda_j \neq 0 \, \forall j$$

*And further, the **span** or **rank** of A and that of Λ coincide, with*

$$rank(A) \;=\; rank(\Lambda) \;=\; \#\{\lambda_j | j = 1, ..., n; \lambda_j \neq 0\}$$

This lemma also happens to be true for non-diagonalizable matrices, but of course the proof is not nearly as straightforward (see Simon and Blume, Thm. 23.9).

As regards the powers of A, we have:

Lemma 179 *For any diagonalizable $A = V \Lambda V^{-1}$, and any $k \in \mathbb{N}$,*

$$A^k = V \Lambda^k V^{-1}$$

If A is invertible, then

$$A^{-1} = V^{-1} \Lambda^{-1} V$$

and for any $k \in \mathbb{N}$

$$A^{-k} = V^{-1}\Lambda^{-k}V$$

Further, since Λ *is a diagonal matrix with typical element* λ_j*,* Λ^{-1} *and* Λ^k *are diagonal matrices with typical elements* $1/\lambda_i$ *and* λ_j^k*, respectively.*

The last result suggests that for diagonalizable matrices A we may generalize the definition of A^k for any non-integer $k \in \mathbb{R}$. Indeed, for any $t \in \mathbb{R}$, and provided $|A| = |\Lambda| \neq 0$ if $t < 0$, we define

$$\Lambda^t \equiv \begin{bmatrix} \lambda_1^t & 0 & ... & 0 \\ 0 & \lambda_2^t & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_n^t \end{bmatrix}$$

and then simply let

$$A^t \equiv V\Lambda^t V^{-1}$$

We emphasize that this definition applies only when A is diagonalizable, and we remind you that any symmetric matrix is diagonalizable.

Exercise 180 *Use the canonical form to prove that for a symmetric matrix* A *it is the case that* $A^r = V'\Lambda^r V$.

Exercise 181 *Use the above result to prove the folowing result:* $\lim_{n->\infty} A^r = 0$*, iff all the eigenvalues of* A *are smaller than 1.*

Exercise 182 *Show that* $|A| = |\Lambda|$.

Exercise 183 *The trace of a square matrix is defined as the sum of its diagonal elements. For instance*

for $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 3 & 1 \end{bmatrix}$, $tr(A) = 4$*. A convenient property of the trace is that it is invariant with respect to cyclical permutations, e.g.* $tr(ABC) = tr(CAB) = tr(BCA)$*. Use this result to prove that for a symmetric matrix* $tr(A) = tr(\Lambda)$*. Note: in fact, this is true for nondiagonalizable matrices as well, but once again the proof is not as straightforward or instructive.*

Exercise 184 *For many applications one is interested in finding a matrix* P *s.t.:* $P'P = A^{-1}$*, where* A *is some symmetric matrix. Find such a matrix.*

(Hint: The matrix $\Lambda^{-1/2} = \begin{bmatrix} \lambda_1^{-1/2} & 0 & 0 \\ 0 & \lambda_2^{-1/2} & 0 \\ 0 & 0 & \lambda_3^{-1/2} \end{bmatrix}$ *will come in very handy.*

Exercise 185 *Show that* $A^{-1} = V^{-1}\Lambda^{-1}V$*, where* $\Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & 0 \\ 0 & \lambda_2^{-1} & 0 \\ 0 & 0 & \lambda_3^{-1} \end{bmatrix}$

Exercise 186 *Show that the solutions to* $|B - \lambda A| = 0$ *and* $|A^{-1}B - \lambda I| = 0$ *are the same.*

Exercise 187 *Show that, for any diagonalizable* A, $|A^k| = |A|^k$.

In a similar way, for any diagonalizable $A = V \Lambda V^{-1}$, we may let the exponential of A be

$$\exp(A) \equiv V \exp(\Lambda) V^{-1}$$

where $\exp(\Lambda)$ is a diagonal matrix with typical element e^{λ_i}. More generally, we let

$$\exp(At) \equiv V \exp(\Lambda t) V^{-1}$$

where

$$\exp(\Lambda t) = \begin{bmatrix} e^{\lambda_1 t} & 0 & ... & 0 \\ 0 & e^{\lambda_2 t} & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & e^{\lambda_n t} \end{bmatrix}$$

Exercise 188 *For a diagonal matrix* Λ *as above, let* $x(t) = \exp(\Lambda t)$, *which simply means* $x_i(t) = e^{\lambda_i t}$ *for all* i. *Characterize the behavior of* $x(t)$ *as* $t \to \infty$.

Remark: The results about the powers Λ^t and A^t prove useful when we study systems of (discrete-time) difference equations, while $\exp(\Lambda t)$ and $\exp(At)$ when we examine systems of (continuous-time) differential equations.

Lemma 189 *Consider the linear system of ordinary differential equations* $\dot{x} = Ax$. *A solution to that is* $x(t) = \exp(At)x(0)$.

Lemma 190 *Consider the linear system of difference equations* $x_{t+1} = Ax_t$. *A solution to that is* $x_t = A^t x_0$.

7.3 Quadratic Forms

We define a **quadratic form** as a function $Q : \mathbb{R}^n \to \mathbb{R}$ such that

$$Q(x) = x'Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

Without loss of generality, we may assume $a_{ij} = a_{ji}$ and thus A to be symmetric. Note that $Q(x)$ is a scalar.

Exercise 191 *Is the assumption that* A *is symmetric really 'without loss of generality?' Take* $Q(x) = x'Ax$ *for arbitrary* A *and show that there exists symmetric* B *such that* $Q(x) = x'Ax = x'Bx$ *(but see Theorem 193 for a hint).*

Definition 192 *A quadratic form* Q *is positive (negative) semidefinite iff* $Q(x) \geq 0$ (≤ 0) *for all* x. *It is positive (negative) definite iff* $Q(x) > 0$ (< 0) *for all* $x \neq 0$.

Note that postive (negative) definiteness implies positive (negative) semidefiniteness, but not the converse. If a quadratic form satisfies none of these conditions, we say it is **indefinite**.

In many economic applications (e.g., static or dynamic optimization, econometrics, etc.), it is important to determine whether a symmetric matrix is positive/negative definite/semidefinite. The diagonalization of the symmetric matrix A can indeed help us easily characterize the quadratic form $Q(x) = x'Ax$.

Since A is symmetric, it is necessarily diagonalizable. Letting V be the orthonormal matrix of eigenvectors and Λ the diagonal matrix of eigenvalues, we have $V' = V^{-1}$ and $V'AV = \Lambda$, or $A = V\Lambda V'$. Hence

$$\begin{aligned} Q(x) &= x'Ax \\ &= x'V\Lambda V'x \\ &= (V'x)\Lambda(V'x) \\ &= \tilde{x}'\Lambda\tilde{x} \;\equiv\; R(\tilde{x}) \end{aligned}$$

where $\tilde{x} = V'x = V^{-1}x$.

Once again, $x \mapsto \tilde{x} = V^{-1}x$ is simply a change of basis, and the quadratic forms Q and R are equivalent. This means that the properties of R are inherited to Q, and vice versa. In particular,

$$Q(x) \lesseqqgtr 0 \ \forall x \quad \Leftrightarrow \quad R(\tilde{x}) \lesseqqgtr 0 \ \forall \tilde{x}$$

Therefore, Q will be positive/negative definite/semidefinite if and only if so is R.

Now notice that, since Λ is a diagonal matrix with diagonal element λ_i, the quadratic form R is a simple sum (a cone indeed) over all eigenvalues λ_i:

$$R(\tilde{x}) \equiv \tilde{x}'\Lambda\tilde{x} = \sum_{i=1}^{n} \lambda_i \tilde{x}_i^2$$

Since $\tilde{x}_i^2 \geq 0$ for all $\tilde{x}$ (or all x), the following theorem is immediate:

Theorem 193 *Let B be an $n \times n$ matrix, let $Q(x) = x'Ax$ be the corresponding quadratic form with $A = \frac{B+B'}{2}$ symmetric***, and let $\{\lambda_i\}_{i=1}^n$ be the eigenvalues of A (possibly not all distinct). Then:*
(i) A is ***positive definite*** *if and only if all eigenvalues are positive,*

$$Q(x) > 0 \ \ \forall \ x \neq 0 \quad \Leftrightarrow \quad \lambda_i > 0 \, \forall i$$

(ii) A is ***negative definite*** *if and only if all eigenvalues are negative,*

$$Q(x) < 0 \ \ \forall \ x \neq 0 \quad \Leftrightarrow \quad \lambda_i < 0 \, \forall i$$

(iii) A is ***positive semidefinite*** *if and only if all eigenvalues are nonnegative,*

$$Q(x) \geq 0 \ \forall x \ \Leftrightarrow \ \lambda_i \geq 0 \, \forall i$$

(iv) A is ***negative semidefinite*** *if and only if all eigenvalues are nonnegative,*

$$Q(x) \leq 0 \;\forall x \;\Leftrightarrow\; \lambda_i \leq 0 \,\forall i$$

and finally
(v) A is indefinite if and only if there are eigenvalues with opposite signs.

***Note that this theorem must make use of the eigenvalues of A, not B! For one thing, we cannot guarantee that B is diagonalizable, but we know this is true of A. Moreover, consider the following:

Question: for a $n \times n$ matrix A (not necessarily symmetric) to be positive definite (in the sense that $x'Ax > 0$ for any nonzero $x \in \mathbb{R}^n$), is it necessary and/or sufficient that its real eigenvalues are all positive?

Answer: It is necessary. Indeed, assume that $\lambda < 0$ is an eigenvalue of A and v is an eigenvector for this eigenvalue. Then $v'Av = \lambda v'v = \lambda |v|^2 < 0$, so A is not positive definite.

On the other hand, it is not sufficient. Consider $A = \begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix}$. Its only eigenvalue is 1, but for $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ we have $x'Ax = -3$.

However, positive definiteness is inherently a property of a quadratic form, not of a matrix (although it can be defined, as above, in terms of a matrix). Remember that there exists infinitely many matrices representing a particular quadratic form (that is, such matrices A that $Q(x) = x'Ax$), all with generally different eigenvalues, and exactly one of them is symmetric. What you want to do to establish positive definiteness (or lack thereof) of a quadratic form is to find this symmetric matrix representing it (if you have any matrix B then $\frac{B+B'}{2}$ is what you are looking for) and test whether its eigenvalues are all positive either by finding them all or by applying the principal minor method or otherwise).

For example, the symmetric matrix representing the same quadratic form as $\begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix}$ is $\begin{pmatrix} 1 & -2.5 \\ -2.5 & 1 \end{pmatrix}$; its determinant is negative, so clearly it does not have both eigenvalues positive and hence the quadratic form is not positive definite, as I demonstrated explicitly above.

Exercise 194 *Let X be an arbitrary real matrix. Show that $X'X$ is positive semidefinite.*

Theorem 195 ***Exercise 196*** *Let X be an $m \times n$ matrix with $m \geq n$ and $rk(X) = n$. Show that $X'X$ is positive definite.*

Exercise 197 *Show that a positive definite matrix is nonsingular*

Exercise 198 *Show that if X is symmetric and idempotent, then X is also positive semi-definite.*

7.4 Static Optimization and Quadratic Forms

The definiteness of a symmetric matrix plays an important role in economic theory. From single-variable calculus, we are familiar with the idea that the sign of the second derivative $f''(x_0)$ of a function $f(x)$ at a critical point x_0 gives (assuming the second derivative exists) a necessary and sufficient condition for determining whether x_0 is a maximum of f, a minimum of f, or neither. The generalization of this test to higher dimensions involves checking the definiteness of the symmetric matrix of cross-partial second derivatives (or Hessian) of f. And fortunately, the intuition learned from single-variable calculus carries over into higher dimensions: just as $f''(x_0) \leq 0$ is necessary and sufficient for x_0 to be a (local) maximizand of twice-differentiable f, so we find that a function of more than one variable is concave at a critical point if the matrix of second derivatives evaluated at that point is negative definite.

Example 199 *Consider the function $Q(x, y) = 2x^3 + xy^2 + 5x^2 + y^2$. The first order conditions with respect to x and y give*

$$\begin{aligned} 6x^2 + 10x + y^2 &= 0 \\ 2(x+1)y &= 0 \end{aligned}$$

These conditions give the following four critical points: $(0, 0), (-\frac{5}{3}, 0), (-1, 2), (-1, -2)$.

The Hessian is $H(x, y) = \begin{bmatrix} 12x + 10 & 2y \\ 2y & 2(x+1) \end{bmatrix}$. *Then we have*

$H(0, 0) = \begin{bmatrix} 10 & 0 \\ 0 & 2 \end{bmatrix}$. *The quadratic form $x'H(0, 0)x = 10x^2 + 2y^2 > 0$ for all $x \neq 0$, so this is a local minimum. Similarly we can find $x'H(-\frac{5}{3}, 0)x = -10x^2 - \frac{4}{3}y^2 < 0$ for all $x \neq 0$, so this is a local maximum. However, $x'H(-1, 2)x = 8xy - 2x^2$, and $x'H(-1, -2)x = -8xy - 2x^2$, both of which are of ambiguous sign. These quadratic forms are indefinite, so these critical points are neither maxima nor minima.*

7.5 Envelope Theorem

When there is a parameter in the optimization problem, how does the value function (the value of f at the optimum) depend on it? Let's start with the simplest case: Unconstrained optimization:

Theorem 200 *$f : U \times I \to R$ where $U \subset R^n$ open and $I \subset R$ interval is C^1 :*

$$f(x, q)$$

Suppose that for each q, there is a solution $x^*(q)$. If $V(q) = f(x^*(q), q) = \max_{x \in R^n} f(x, q)$, Suppose that $q \to x^*(q)$ is of class C^1, then:

$$\frac{dV(q)}{dq} = \frac{\partial f}{\partial q}(x^*(q), q)$$

Proof. Take the first order condition: $\frac{\partial f}{\partial x}(x^*(q), q) = 0$. Now, $V'(q) = \frac{\partial f}{\partial q}(x^*(q), q) + x^{*\prime}(q)\frac{\partial f}{\partial x}(x^*(q), q) = \frac{\partial f}{\partial q}(x^*(q), q)$ ■

We can generalize this:

Theorem 201 *Generalization:*

Let $f(x;q)$ be a continuous function, and $x^*(q)$ denote the solution to the problem of maximizing $f(x;q)$ on the constraint set $h_i(x;q) = 0$, $i = 1, ..., k$. The Lagrangian is $\Lambda = f(x;q) - \lambda h(x;q)$ (where the constraints are written as a vector). Then

$$\begin{aligned} \frac{df(x^*(q);q)}{dq} &= \left.\frac{\partial \Lambda}{\partial q}\right|_{x=x^*(q)} \\ &= \left.\frac{\partial f(x;q)}{\partial q}\right|_{x=x^*(q)} - \lambda \left.\frac{\partial h(x;q)}{\partial q}\right|_{x=x^*(q)} \end{aligned}$$

Like in the unconstrained case, this says that we can "ignore" the effect of x changing as q changes, and focus on the direct effect of changes in q.

Proof. *Define the 'value function' as the value of the objective function at the maximum. The value function is written as a function of q, not x, because it is assumed that for any given q, the x's are simply whichever maximize the objective function for the given q (this is why we write $x^*(q)$). Denote this as*

$$M(q) = f(x^*(q), q)$$

Then differentiating both sides with respect to q, we have

$$\frac{dM(q)}{dq} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial q} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial f}{\partial q}$$

From the Lagrangian we have

$$\frac{\partial f}{\partial x_i} = \lambda \frac{\partial h}{\partial x_i}$$

Substituting, we have

$$\frac{dM(q)}{dq} = \lambda\frac{\partial h}{\partial x_1}\frac{\partial x_1}{\partial q} + \lambda\frac{\partial h}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \lambda\frac{\partial h}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial f}{\partial q}$$

Since we know that at the optimum, the constraints bind, we can differentiate the following identity:

$$\begin{aligned} h(x;q) &= 0 \\ \frac{\partial h}{\partial x_1}\frac{\partial x_1}{\partial q} + \frac{\partial h}{\partial x_2}\frac{\partial x_2}{\partial q} + ... + \frac{\partial h}{\partial x_n}\frac{\partial x_n}{\partial q} + \frac{\partial h}{\partial q} &= 0 \end{aligned}$$

Substituting again, we have the desired result:

$$\frac{dM(q)}{dq} = -\lambda\frac{\partial h}{\partial q} + \frac{\partial f}{\partial q}$$

■

Application: Consider maximization subject to budget constraints

$$\mathbf{V}(I) = \max_x U(x)$$
$$px \le I$$

In this example

$$\frac{d\mathbf{V}(I)}{dI} = \lambda,$$

so λ is the marginal utility of wealth (called also the shadow price of the constraint).