Solutions to Introductory Exam for September, 2004

1. Prove that the intersection $\mathcal{O}_1 \cap \mathcal{O}_2$ of two open sets $\mathcal{O}_1$ and $\mathcal{O}_2$ in $\mathbb{R}^n$ is open.

Solution:

Let $x \in \mathcal{O}_1 \cap \mathcal{O}_2$. Since $x \in \mathcal{O}_1$ and $\mathcal{O}_1$ is open, there is a $\delta_1 > 0$ such that all y with $|x-y| < \delta_1$ belong to $\mathcal{O}_1$. Similarly, there is a $\delta_2 > 0$ such that all y with $|x-y| < \delta_2$ belong to $\mathcal{O}_2$. Take δ to be the smaller of δ_1 and δ_2. Then $\delta > 0$, and if $|x-y| < \delta$, then y belongs to both $\mathcal{O}_1$ and $\mathcal{O}_2$, i.e., y belongs to $\mathcal{O}_1 \cap \mathcal{O}_2$.

2. Suppose n persons arrive at a party, each bearing a gift. The gifts are collected and distributed at random to the n guests. What is the average number of persons who receive the same gift they brought to the party?

Solution:

Let $X_i = 1$ if person i receives his own gift and $X_i = 0$ otherwise, $i = 1, \ldots, n$. Then $P(X_i = 1) = \frac{1}{n}$ for each i and $N = \sum_{i=1}^n X_i$ is the number of persons who receive their own gift. But

$$\mathbb{E}\, N = \mathbb{E} \sum_{i=1}^n X_i = \sum_{i=1}^n \mathbb{E}\, X_i = \sum_{i=1}^n \tfrac{1}{n} = 1.$$

3. Let $u = [u_1, u_2, u_3]^T$ and $v = [v_1, v_2, v_3]^T$ be vectors in $\mathbb{R}^3$.

(a) Show that the cross product $u \times v$ is orthogonal to u.

(b) Show that *Lagrange's identity*

$$\|u \times v\|^2 = \|u\|^2 \|v\|^2 - (u \cdot v)^2$$

holds.

Solution:

(a) We have

$$\begin{aligned} u \cdot (u \times v) &= [u_1, u_2, u_3]^T \cdot [u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1]^T \\ &= u_1(u_2 v_3 - u_3 v_2) + u_2(u_3 v_1 - u_1 v_3) + u_3(u_1 v_2 - u_2 v_1) = 0. \end{aligned}$$

(b) Since

$$\|u \times v\|^2 = (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2$$

and

$$\|u\|^2\|v\|^2 - (u \cdot v)^2 = (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2,$$

Lagrange's identity can be established by "multiplying out" the right sides of these two equations and verifying their equality.

4. Let $x, y \in \mathbb{R}^n$. Prove that

$$\left|\|x\|^2 - \|y\|^2\right| \le \left[\sum_j (x_j + y_j)^2 \cdot \sum_j (x_j - y_j)^2\right]^{1/2}.$$

The sums are over all j from 1 through n.

Solution:

By the Cauchy–Schwarz inequality,

$$|(x + y) \cdot (x - y)| \le \|x + y\|\|x - y\|.$$

The dot product on the left hand side of this last inequality simplifies as follows:

$$(x + y) \cdot (x - y) = x \cdot x + y \cdot x - x \cdot y - y \cdot y = \|x\|^2 - \|y\|^2,$$

and the right hand side is exactly

$$\left[\sum_j (x_j + y_j)^2 \cdot \sum_j (x_j - y_j)^2\right]^{1/2},$$

giving the result.

5. Prove that if $E_1, E_2, E_3, \ldots, E_n$ are independent events, then

$$\mathbb{P}(\cup_{i=1}^n E_i) = 1 - \prod_{i=1}^n [1 - \mathbb{P}(E_i)].$$

Solution:

By DeMorgan's Law,

$$(\cup_{i=1}^n E_i)^c = \cap_{i=1}^n E_i^c,$$

so

$$\mathbb{P}\left(\cup_{i=1}^{n} E_i\right) = 1 - \mathbb{P}\left(\cap_{i=1}^{n} E_i^c\right).$$

Since the events E_i, $i = 1, 2, \ldots, n$, are independent, their complements are also independent. Thus,

$$\mathbb{P}\left(\cap_{i=1}^{n} E_i^c\right) = \prod_{i=1}^{n} \mathbb{P}(E_i^c),$$

which by complementation is equal to

$$\prod_{i=1}^{n} [1 - \mathbb{P}(E_i)].$$

6. Suppose that the coefficients of the power series $\sum_n a_n x^n$ are integers, infinitely many of which are distinct from zero. Prove that the radius of convergence is at most one.

Solution #1:

A power series is uniformly and absolutely convergent on any compact subinterval of its interval of convergence. But $\sum_n |a_n| = \infty$, so $R \leq 1$.

Solution #2:

For each nonzero integer a_n, the nth root $|a_n|^{1/n}$ is at least 1. By assumption, there are infinitely many such coefficients. Hence, by the root test the radius of convergence R satisfies

$$\frac{1}{R} = \limsup_{n \to \infty} \sqrt[n]{|a_n|} \geq 1,$$

or $R \leq 1$.

7. A trucker has to transport an unstable cargo, in rough terrain, from an origin O to a destination D over a network of one-way road-links. Associated with each link L is a calculable probability $p(L)$ that the cargo will explode while being transported over L. Consider the problem of finding a route, from O to D, that maximizes the probability that the cargo will complete the trip without exploding.

(a) **(4 points)** Show how, under an appropriate assumption of independence, this can be transformed into a standard shortest-path problem solvable by the usual methods for such problems.

(b) **(1 point)** Under what (physical) circumstances might the assumption of independence be unreasonable?

Solution:

(a) We use the notation $q(L) = 1 - p(L)$. Consider any path P from O to D, composed of the successive links $L_1, L_2, \ldots, L_n$. If this route is used, the probability that the cargo traverses link L_i safely (*if* it gets that far without exploding) is $q(L_i)$. Thus (with a final remark to follow) the probability that the cargo will survive the whole trip along route P is $q(L_1)q(L_2)\cdots q(L_n)$. So we want to choose P to maximize this. The usual shortest-path problem and its solution methods call for minimizing (not maximizing) a sum (not a product) of the nonnegative "lengths" of the links in the route. But the stated problem is equivalent to minimizing $(-\log[q(L_1)]) + (-\log[q(L_2)) + \cdots + (-\log[q(L_n)])$.

So we can reach the standard form by taking the "length" of each link L to be $(-\log[q(L)])$. And since the q's are probabilities, these "lengths" are indeed nonnegative. [There is trouble if $q(L) = 0$, i.e., if $p(L) = 1$, but such "sure death" links would be excluded in advance; if that leaves no route at all from O to D, the trip cannot be undertaken.]

(b) The "suspect assumption" is that of independence, implicit in forming the product of the route-links' q-values. In effect this assumes the instability of the cargo is not cumulative, i.e., it ignores the possibility that even if the cargo doesn't explode early in the trip before L_i is reached, the effect of those early shake-ups, if rough enough, may increase its probability of exploding while L_i is being traversed.

8. Let $(a_n : n = 1, 2, \ldots)$ be a sequence of real numbers defined recursively: $a_1 = \sqrt{2}$ and $a_{n+1} = \sqrt{2 + a_n}$ for $n \geq 1$. Show that

$$\lim_{n\to\infty} a_n = A$$

exists and find A.

Solution:

By induction, $a_n \leq 2$ for all $n \geq 1$ since $a_1 \leq 2$ and, if $a_n \leq 2$, then $a_{n+1} = \sqrt{2 + a_n} \leq \sqrt{2 + 2} = 2$. In addition, $a_{n+1} \geq a_n$ for all $n \geq 1$ because $2 + a_n \geq 2a_n \geq a_n^2$, which implies $\sqrt{2 + a_n} \geq a_n$. Since (a_n) is bounded and increasing, it converges. Finally,

$$A = \lim_n a_{n+1} = \lim_n \sqrt{2 + a_n} = \sqrt{2 + A}$$

and consequently $A = 2$.

9. Suppose that a cancer diagnostic test is 95 percent accurate both on those that do and those that do not have the disease. If 0.4 percent of the population have cancer, compute the probability that a tested person chosen randomly from the population has cancer, given that his or her test result indicates so.

Solution: This is a standard Bayes rule problem:

$$\begin{aligned}\mathbb{P}[\text{cancer}|\text{positive}] &= \frac{\mathbb{P}[\text{positive}|\text{cancer}]\,\mathbb{P}[\text{cancer}]}{\mathbb{P}[\text{positive}|\text{cancer}]\,\mathbb{P}[\text{cancer}] + \mathbb{P}[\text{positive}|\text{cancer}^c]\,\mathbb{P}[\text{cancer}^c]} \\ &= \frac{(0.95)(0.004)}{(0.95)(0.004)+(0.05)(0.996)} = \frac{0.0038}{0.0536} = \frac{19}{268} \approx 0.0709.\end{aligned}$$

10. Let $D \subseteq \mathbb{R}^p$, and suppose that (f_n) is a sequence of continuous $\mathbb{R}^q$-valued functions that converges uniformly on D to f. Suppose also that (x_n) is a sequence of elements of D that converges to x in D. Prove that $f_n(x_n) \to f(x)$.

Solution:

By the triangle inequality (for the Euclidean norm on $\mathbb{R}^q$),

$$|f_n(x_n) - f(x)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)|\,.$$

But, when n is large, the first term here is small by uniform convergence of f_n to f, and the second term is small because $x_n \to x$ and the limit function f is continuous (because it is the uniform limit of continuous functions on D).

11. For a nonnegative integer k, let $a_k := 2k+1$ and let $b_k := k^2$. Set

$$F(x) := \sum_{k=0}^{\infty} a_k x^k \qquad \text{and} \qquad G(x) := \sum_{k=0}^{\infty} b_k x^k.$$

Prove that

$$G(x) = \frac{x}{1-x} F(x).$$

[Note: You need not discuss the convergence of the various series encountered.]

Solution #1:

The desired result is equivalent to $(1-x)G(x) = xF(x)$, i.e., to $b_0 = 0$ and $b_k - b_{k-1} = a_{k-1}$ for $k \geq 1$. Indeed, $b_0 = 0^2 = 0$, and for $k \geq 1$ we have $b_k - b_{k-1} = k^2 - (k-1)^2 = 2k - 1 = a_{k-1}$.

Solution #2:

It is well known (and easy to prove by induction) that the sum of the first n positive odd numbers is n^2. In terms of the a's and b's, this can be expressed as

$$b_n = a_0 + a_1 + \cdots + a_{n-1}. \tag{$*$}$$

Let c_k be defined by $c_0 = 0$ and $c_k = 1$ for all $k \geq 1$. Then $(*)$ can be rewritten

$$b_n = a_0 c_n + a_1 c_{n-1} + \cdots + a_{n-1} c_1 + a_n c_0. \tag{$**$}$$

Let $H(x)$ be the ordinary generating function for the c's, i.e., $H(x) = \sum_{k=0}^{\infty} c_k x^k$. Then $(**)$ implies

$$G(X) = H(X)F(X)$$

by the convolution property. Finally, $H(x)$ is a geometric series

$$H(x) = x + x^2 + x^3 + x^4 + \cdots = \frac{x}{1-x},$$

and the result follows.

[Alternatively we could work out the expressions for F and G directly, but that's more work and boring.]

12. If X and Y are real-valued random variables with joint density $f_{X,Y}$, show that the density function f_Z of the product $Z = XY$ is given by the improper integral

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x, z/x)|x|^{-1}\, dx.$$

Solution:

Let T map (x, y) into (z, w) by

$$z = xy, \quad w = x.$$

The inverse T^{-1} maps (z, w) into (x, y) by

$$x = w, \quad y = z/w,$$

and the Jacobian is

$$\frac{\partial(x, y)}{\partial(z, w)} = -w^{-1}.$$

Thus, $f_{Z,W}(z, w) = f_{X,Y}(w, z/w)|w|^{-1}$. Integrate over w to conclude the argument.

13. Let

$$A = \begin{bmatrix} 1 & \rho \\ \rho & \rho^2 \end{bmatrix},$$

where $\rho \in \mathbb{R}$.

(a) Compute the eigenvalues and eigenvectors of this matrix.

(b) For what values of ρ is A positive definite?

(c) For what values of ρ is A positive semidefinite?

Solution:

(a) The characteristic polynomial of A is

$$\det \begin{bmatrix} t-1 & -\rho \\ -\rho & t-\rho^2 \end{bmatrix} = t^2 - (1+\rho^2)t,$$

whose roots are the eigenvalues $\lambda = 0$ and $\lambda = 1+\rho^2$. Solving $A\,[x,y]^T = \lambda\,[x,y]^T$ for each eigenvalue leads to the following:

- For $\lambda = 0$: $x + \rho y = 0$, so $[-\rho, 1]^T$ generates the one-dimensional eigenspace corresponding to this eigenvalue.
- For $\lambda = 1+\rho^2$: $y = \rho x$ so $[1, \rho]^T$ generates the one-dimensional eigenspace corresponding to this eigenvalue.

(b) Since 0 is an eigenvalue of A no matter what ρ is, the matrix is never positive definite (it's not even invertible).

(c) Since the matrix is real and symmetric and has nonnegative eigenvalues it must be positive semidefinite.

14. Given real numbers x and y, the so-called n-dimensional *combinatorial matrix* A is defined as the $n \times n$ matrix

$$A := xI + yJ,$$

where I is the identity matrix and J is the matrix of all 1's.

(a) Compute the determinant of A.

(b) Show that A is invertible if and only if $x \neq 0$ and $x + ny \neq 0$, and that

$$A^{-1} = \frac{1}{x(x+ny)}[(x+ny)I - yJ]$$

in that case.

Solution:

(a) Subtract column 1 from each of columns $2, \ldots, n$. Then add each of rows $2, \ldots, n$ to row 1. The result is a lower-triangular determinant with diagonal $(x+ny, x, \ldots, x)$. So $\det A = x^{n-1}(x+ny)$.

(b) The necessary and sufficient condition for invertibility is immediate from the solution to part (a). If the condition holds, denote the alleged inverse of A by B. Then

$$AB = \frac{1}{x(x+ny)}(xI+yJ)[(x+ny)I-yJ] = \frac{1}{x(x+ny)}[x(x+ny)I+ny^2J-y^2J^2] = I,$$

as desired, since $J^2 = nJ$.

15. Let $n \geq 3$ be an integer and let $N = \{1, 2, \ldots, n\}$. Form a graph G in which $V(G)$ is the set of all 2-element subsets of N and in which (for distinct vertices v and w) $vw \in E(G)$ provided $v \cap w$ is nonempty. Prove that G is Eulerian.

Solution:

Note that for all values of n we have that G is connected. Indeed, for distinct vertices v, w either $v \sim w$ or else $v = \{a, b\}$ and $w = \{c, d\}$ for distinct $a, b, c, d \in N$. In the latter case, $v = \{a, b\} \sim \{a, d\} \sim \{c, d\} = w$. Thus, for any pair of vertices, there is a path that joins them.

And so, G is Eulerian if and only if all vertices in G have even degree. Note that for any vertex $v = \{a, b\}$, we have $d(v) = 2(n-2)$ as there are $n-2$ neighbors of the form $\{a, x\}$ and $n-2$ neighbors of the form $\{b, x\}$. Since $2(n-2)$ is even, G is Eulerian.

16. Suppose that $f : [0, 1] \to [0, 1]$ is such that

$$|f(x) - f(y)| \geq |x - y| \text{ for all } x, y \in [0, 1]. \qquad (*)$$

(a) If f is continuous, show that f is one-to-one and onto.

(b) Find all functions $f : [0, 1] \to [0, 1]$ satisfying the property $(*)$.

Solution:

(a) If $f(x) = f(y)$, then $|x - y| \leq |f(x) - f(y)| = 0$ and $x = y$, so f is one-to-one. Also, $|f(1) - f(0)| \geq 1$ so that $|f(1) - f(0)| = 1$. Since f is continuous, the intermediate value theorem implies that $f([0, 1]) = [0, 1]$ and f is onto.

(b) There are two such functions. Since $|f(1) - f(0)| = 1$, we have (i) $f(0) = 0$ and $f(1) = 1$ or (ii) $f(0) = 1$ and $f(1) = 0$. Consider the first case. It implies both $|f(x)| \geq x$ and $1 - f(x) = |f(x) - 1| \geq |x - 1| = 1 - x$, so that $f(x) = x$. The second case is similar and gives $f(x) = 1 - x$.

17. A fair six-sided die is continually rolled until the sum of the outcomes of all rolls (strictly) exceeds 300. Is the probability that at least 82 rolls are needed smaller or larger than 75%? Explain your answer. [HINT: You may use without proof the fact that the standard deviation of the outcome of a single roll is $\sqrt{35/12} \approx 1.7078$.]

Solution:

The probability in question is the probability that the sum of 81 roll-outcomes does not exceed 300. Using the normal approximation (with continuity correction, whose use is a minor matter here), this probability is approximately the probability that a standard normal does not exceed

$$\frac{300.5 - 81 \times 3.5}{9 \times \sqrt{35/12}} \approx \frac{17}{15.37} \approx 1.1.$$

All students of probability should memorize that the area under the standard normal density between -1 and 1 is about 68%, so the area to the left of 1 is about 84%. Hence the probability in question is **larger than** 75%.

18. LUXO is a company that manufactures luxury cars and trucks. It believes its most likely customers are high-income women and men (HIW's and HIM's, respectively). To reach these groups, it's undertaking an advertising campaign involving the purchase of one-minute commercial spots on two types of TV programs: romantic comedies and football games. Each comedy commercial costs \$50K, and is estimated to be seen by 7 million HIW's and 2 million HIM's; the corresponding figures for "football ads" are: \$100K, 2 million HIW's, and 12 million HIM's. The company would like its commercials to reach at least 28 million HIW's and 24 million HIM's. Initial issue: how many spots of each type (comedy or football) to buy?

(a) Formulate a naive linear program to determine how LUXO should seek to meet its goals at minimum total cost.

(b) Solve the model by whatever method you like, but with a clear explanation.

(c) Identify at least three potential pitfalls in using your model as a valid reliable tool for reaching the company's "what to do?" decision.

Solution:

(a) This is straightforward. After the decision variables

$$C := \text{number of "comedy spots" to buy,}$$
$$F := \text{number of "football spots" to buy}$$

are introduced, the relevant linear program calls for

$$\text{Minimize } Z = 50C + 100F \text{ (units: \$K)}$$

subject to the functional constraints

$$7C + 2F \geq 28 \text{ (units: millions of HIW's)}, \quad (1)$$
$$2C + 12F \geq 24 \text{ (units: millions of HIM's)}, \quad (2)$$

and the non-negativity constraints

$$C \geq 0 \quad (3)$$
$$F \geq 0. \quad (4)$$

(b) The simplex method could be used, but in the presence of only two variables and so few functional constraints, resort to the standard "graphical approach" is natural. In the (C, F)-plane one draws the four lines corresponding to the equation versions (1E–4E) of the constraints, marks the "feasible side" of each of those lines, and identifies (say, by shading or cross-hatching) the feasible region (it's unbounded) as the intersection of the four resultant half-planes. Next, the line $50C + 100F = 0$ is moved parallel to itself, "northeast" (thus generating "increasing" level lines of Z) until it first touches the feasible region. (Alternatively, one could move such a level line "southwest," parallel to itself, until it's about to *leave* the feasible region.) This is readily seen to occur at the intersection of the lines (1E) and (2E), determined by simple algebra to be

$$(C^*, F^*) = (18/5, 7/5) = (3.6, 1.4).$$

So that's the optimal solution; *deduct* some credit if the student fails also to record the resulting value $Z_{\text{min}} = \$320\text{K}$.

(c) By its nature, part (c) lacks a definite "right answer," but a thoughtful reply would presumably address some of the following points (and possibly other good ones as well):

(i) The "penetration coefficients" in (1)–(2) are "estimates," presumably based on data (Nielsen ratings?) gathered, with some sampling error, for products and TV programs thought analogous to those here. So these "data" are not reliably certain and precise.
Also, "football games" differ in attractiveness (e.g., "bowl" vs. "exhibition"), as do romantic comedies and even specific episodes (like a much-hyped end-of-season cliff-hanger).

(ii) Some people watch both romantic comedies and football games. The formulation of (1)–(2) double-counts such viewers. Some people watch *several* football games, or several comedies. Should we be counting *viewings* or *viewers*? And how is either count transformable into an estimate of additional product-sales, the real aim of the advertising? Repeated ad-viewings may be reinforcing up to a point, off-putting beyond it.

(iii) *Are* fractional answers OK? If one is buying "total minutes" rather than some "number of one-minute spots," fractions could be OK, but that's a deviation from the exact wording above.

(iv) The linearity of Z rules out considering such possibilities as "volume discounts."

(v) Just how is "high-income" defined? Is the definition the same in all parts of the country? Is it the same for women as for men? *Should* it be? For luxury vehicles, is the important attribute "income" or "wealth"?

(vi) What about the advertising of LUXO's competitors? Also, at some point the analysis will need to get specific about the diversity, frequency, and exact selection of vehicles (games or comedy shows) for the ads. What about the *costs* of developing the commercials, as distinct from that of purchasing the "network time" in which to run them?

19. The *convex hull* of a set $X \subseteq \mathbb{R}^n$ is defined as the set of points of the form $\sum_{i=1}^k \lambda_i x_i$ where, for each i, we have $x_i \in X$ and $\lambda_i \geq 0$, and where $\sum_{i=1}^k \lambda_i = 1$. Denote the convex hull of X by $\mathrm{conv}(X)$. Let $X \subseteq \mathbb{R}^n$ be a finite set, and let $C := \mathrm{conv}(X)$.

(a) Prove that C is closed under the formation of convex combinations, that is, if $x, y \in C$, then $\lambda x + (1 - \lambda)y \in C$ for all $\lambda \in [0, 1]$.

(b) Prove that C is a closed set.

Solution:

(a) Let $X = \{x_1, \ldots, x_m\}$, and observe that since we can set any coefficients to zero that do not appear, a member of C can always be represented in the form

$$\sum_{i=1}^{m} \lambda_i x_i$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^m \lambda_i = 1$. Now if $x, y \in C$, then we can write

$$x = \sum_{i=1}^{m} \lambda_i x_i$$

where $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$, and

$$y = \sum_{i=1}^{m} \omega_i x_i$$

where $\omega_i \geq 0$ and $\sum_i \omega_i = 1$.

If $\lambda \in [0,1]$ we have

$$\lambda x + (1-\lambda)y = \lambda \sum_{i=1}^{m} \lambda_i x_i + (1-\lambda)\sum_{i=1}^{m} \omega_i x_i$$
$$= \sum_{i=1}^{m} \{\lambda\lambda_i + (1-\lambda)\omega_i\} x_i,$$

and we have $\lambda\lambda_i + (1-\lambda)\omega_i \geq 0$ and

$$\sum_{i=1}^{m} \{\lambda\lambda_i + (1-\lambda)\omega_i\} = \lambda \sum_{i=1}^{m} \lambda_i + (1-\lambda)\sum_{i=1}^{m} \omega_i = \lambda + (1-\lambda) = 1.$$

(b) To see that C is a closed set, suppose we have a sequence $x^{(N)} \subseteq C$ that converges to a limit $x \in \mathbb{R}^n$. Then we can write

$$x^{(N)} = \sum_{i=1}^{m} \lambda_i^{(N)} x_i$$

where $\lambda_i^{(N)} \geq 0$ and $\sum_i \lambda_i^{(N)} = 1$. Since the simplex

$$\{(\lambda_1, \ldots, \lambda_m) : \lambda_i \geq 0,\ \sum_i \lambda_i = 1\}$$

is compact, we can find a subsequence (N_p) for which the vectors $\lambda^{(N_p)}$ converge to λ coordinatewise, with λ in the simplex. Then we have

$$x^{(N_p)} = \sum_{i=1}^{m} \lambda_i^{(N_p)} x_i \to \sum_{i=1}^{m} \lambda_i x_i = x,$$

and we conclude that $x \in \mathrm{conv}(X)$.

20. Use the spectral theorem for positive semidefinite real symmetric matrices (or more generally for real symmetric matrices) to prove that if $A = [a_{ij}]$ and $B = [b_{ij}]$ are $n \times n$ positive semidefinite real symmetric matrices, then so is $C = [a_{ij}b_{ij}]$. [HINT: Determine what the spectral theorem gives for the form of the entries a_{ij}.]

Solution:

By the spectral theorem, we can write $a_{ij} = \sum_{k=1}^{n} \lambda_k u_{ik} u_{jk}$ with $\lambda_k \geq 0$ for every k, where the vectors $u_i = [u_{ik}]$ $(i = 1, \ldots, n)$ are real orthonormal eigenvectors of A. Choose any $x \in \mathbb{R}^n$. Then

$$x^T C x = \sum_{i,j=1}^{n} c_{ij} x_i x_j = \sum_{k=1}^{n} \lambda_k \sum_{i,j} b_{ij}(x_i u_{ik})(x_j u_{jk}) \geq 0,$$

since the sum $\sum_{i,j}$ for each fixed k is nonnegative because B is positive semidefinite.