


 1 Show that each A_k [$A(1 : k, 1 : k)$ in matlab notation] is SPD.


Solution: Let x be any vector in \mathbb{R}^k and consider the vector y of \mathbb{R}^n obtained by stacking x followed by $n - k$ zeros. Then it can be easily seen that : $(A_k x, x) = (A y, y)$ and since A is SPD then $(A y, y) > 0$ and therefore $(A_k x, x) > 0$ for any x in \mathbb{R}^k . Hence A_k is SPD. \square

 2 Consequence $\det(A_k) > 0$

Solution: This is because the determinant is the product of the eigenvalues which are real positive (see notes). \square

 3 If A is SPD then for any $n \times k$ matrix X of rank k , the matrix $X^T A X$ is SPD.

Solution: For any $v \in \mathbb{R}^k$ we have $(X^T A X v, v) = (A X v, X v)$. In addition, since X is of full rank, then $X v$ cannot be zero if v is nonzero. Therefore we have $(A X v, X v) > 0$. \square

 4 Show that if $A^T = A$ and $(A x, x) = 0 \forall x$ then $A = 0$.

Solution: The condition implies that for all x, y we have $(A(x + y), x + y) = 0$. Now expand this as: $(Ax, x) + (Ay, y) + 2(Ax, y) = 0$ for all x, y which shows that $(Ax, y) = 0 \forall x, y$. This implies that $A = 0$ (e.g. take $x = e_j, y = e_i$)... \square

5 Show: A nonzero matrix A is indefinite iff:

$$\exists x, y : (Ax, x)(Ay, y) < 0.$$

Solution:

\leftarrow Trivial. The matrix cant be PSD or NSD under the conditon

\rightarrow Need to prove: If A is indefinite then there exist such that $x, y : (Ax, x)(Ay, y) < 0$. Assume contrary is true, i.e.,

$$\forall x, y (Ax, x)(Ay, y) \geq 0$$

. There is at least one x_0 such that (Ax_0, x_0) is nonzero, otherwise $A = 0$ from previous question. Assume $(Ax_0, x_0) > 0$. Then $\forall y (Ax_0, x_0)(Ay, y) \geq 0$. which implies $\forall y : (Ay, y) \geq 0$, i.e., A is positive semi-definite. This contradicts the assumption that A is neither positive nor negative semi-definite \square

6 The (standard) LU factorization of an SPD matrix A exists.

Solution: This is an immediate consequence of the main theorem on existence (Lec. notes. set #5) and Exercise 1 in this set which showed that $\det(A_k) > 0$ for $k = 1, \dots, n$. \square

Example:

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -1 & 5 & 0 \\ 2 & 0 & 9 \end{pmatrix}$$

 7 Is A symmetric positive definite?

Solution: Answer is yes because $\det(A_k) > 0$ for $k = 1, 2, 3$. \square

 8 What is the LDL^T factorization of A ?

Solution: The LU factorization is:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1/2 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{pmatrix}$$

Therefore $A = LDL^T$ where L is as given above and

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \quad \square$$

 9 What is the Cholesky factorization of A ?

Solution: From the above LDLT factorization we have $A = GG^T$

with

$$G = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & 1 & 2 \end{pmatrix} \quad \square$$

Gradient of $\psi(\mathbf{x}) = (\mathbf{A}\mathbf{x}, \mathbf{x})$

In practice exercise # 6 it is asked: Let \mathbf{A} be symmetric and $\psi(\mathbf{x}) = (\mathbf{A}\mathbf{x}, \mathbf{x})$. What is the partial derivative $\frac{\partial\psi(\mathbf{x})}{\partial x_k}$? What is the gradient of ψ ?

Solution: First note that

$$\psi(\mathbf{x}) = \sum_{i=1}^n x_i \left[\sum_{j=1}^n a_{ij} x_j \right]$$

and so, using basic rules for derivatives of products:

$$\begin{aligned} \frac{\partial\psi(\mathbf{x})}{\partial x_k} &= \sum_{i=1}^n \frac{\partial x_i}{\partial x_k} \left[\sum_{j=1}^n a_{ij} x_j \right] + \sum_{i=1}^n x_i \left[\frac{\partial x_i}{\partial x_k} \sum_{j=1}^n a_{ij} x_j \right] \\ &= \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n x_i a_{ik} \\ &= 2 \sum_{j=1}^n a_{kj} x_j \end{aligned}$$

which is nothing but twice the k -th component of $\mathbf{A}\mathbf{x}$ or $\frac{\partial\psi(\mathbf{x})}{\partial x_k} = 2(\mathbf{A}\mathbf{x})_k$. Therefore the gradient of ψ is

$$\nabla\psi(\mathbf{x}) = 2\mathbf{A}\mathbf{x}.$$

A somewhat simpler solution for finding the gradient is to expand $\psi(\mathbf{x} + \boldsymbol{\delta}) = (\mathbf{A}(\mathbf{x} + \boldsymbol{\delta}), (\mathbf{x} + \boldsymbol{\delta})) = \dots$ and to write that the

linear term should be of the form $[\nabla\psi]^T \delta$.