Positive definiteness and semidefiniteness

Zdeněk Dvořák

May 19, 2015

For integers a, b, and c, let $D(a,b,c)$ be the diagonal matrix with

$$D_{i,i} = \begin{cases} +1 & \text{for } i = 1,\ldots,a, \\ -1 & \text{for } i = a+1,\ldots,a+b, \\ 0 & \text{for } i = a+b+1,\ldots,a+b+c. \end{cases}.$$

Definition 1. *We say that a quadratic form f on a vector space $\mathbf{V}$ of finite dimension has signature (a,b,c) if there exists a basis B such that $[f]_B = D(a,b,c)$.*

A symmetric matrix A has signature (a,b,c) if the quadratic form $f(x) = x^T Ax$ has signature (a,b,c); i.e., $A = BD(a,b,c)B^T$ for some regular matrix B.

Lemma 1. *Let A be a real symmetric matrix. If f has signature (a,b,c), then the sum of algebraic multiplicities of the positive eigenvalues of A is a, and the sum of algebraic multiplicities of the negative eigenvalues of A is b.*

1 Positive definiteness and semidefiniteness

Definition 2. *Let $\mathbf{V}$ be a vector space over real numbers. A symmetric bilinear form $b : \mathbf{V} \times \mathbf{V} \to \mathbf{R}$ is positive semidefinite if $b(x,x) \ge 0$ for all $x \in \mathbf{V}$, and it is positive definite if $b(x,x) > 0$ for all $x \in \mathbf{V} \setminus \{o\}$.*

Similarly, a quadratic form $f : \mathbf{V} \to \mathbf{R}$ is positive semidefinite if $f(x) \ge 0$ for all $x \in \mathbf{V}$, and it is positive definite if $f(x) > 0$ for all $x \in \mathbf{V} \setminus \{o\}$.

A real symmetric $n \times n$ matrix A is positive semidefinite if $x^T Ax \ge 0$ for all $x \in \mathbf{R}^n$, and it is positive definite if $x^T Ax > 0$ for all $x \in \mathbf{R}^n \setminus \{o\}$.

We write $X \succeq 0$ if X is positive semidefinite, and $X \succ 0$ if X is positive definite.

Observation 2. *Let $\mathbf{V}$ be a vector space over real numbers of finite dimension and let C be its basis. A symmetric bilinear form $b : \mathbf{V} \times \mathbf{V} \to \mathbf{R}$ is positive (semi)definite if and only if the quadratic form $f(x) = b(x, x)$ is positive (semi)definite, and if and only if the matrix $[b]_C$ is positive (semi)definite.*

Observation 3. *A symmetric bilinear form defines an inner product if and only if it is positive definite. Conversely, every inner product is a symmetric bilinear form.*

Lemma 4. *Let A be a symmetric real $n \times n$ matrix. If $A \succ 0$, then A is regular, and $A^{-1} \succ 0$.*

Proof. If $Ax = o$, then $x^T Ax = 0$, and thus $x = o$ since A is positive definite; hence, A is regular. For any $x \in \mathbf{R}^n \setminus \{o\}$, we have $x^T A^{-1} x = x^T (A^{-1}A)A^{-1}x = (A^{-1}x)^T A(A^{-1}x) > 0$. □

Lemma 5. *Let A be a symmetric real $n \times n$ matrix. The following claims are equivalent.*

1. *$A \succeq 0$.*
2. *A has signature $(p, 0, z)$ for some integers $p, z \geq 0$ such that $p + z = n$.*
3. *All eigenvalues of A are non-negative.*
4. *$A = B^2$ for some symmetric positive semidefinite matrix B.*
5. *$A = B^2$ for some symmetric matrix B.*
6. *$A = F^T F$ for some matrix F.*

Proof. $1 \Rightarrow 2$**:** Let (a, b, c) be the signature of A; then, $D(a, b, c) = B^T AB$ for some regular matrix B, and thus $D(a, b, c)_{i,i} = (Be_i)^T A(Be_i) \geq 0$ for $i = 1, \ldots, n$. Hence, $b = 0$.

$2 \Rightarrow 3$**:** By Lemma 1.

$3 \Rightarrow 4$**:** Since A is symmetric and real, we have $A = Q^T DQ$ for some orthogonal Q and a diagonal matrix D with eigenvalues of A on the diagonal. Hence, all entries of D are non-negative. Let E be the diagonal matrix such that $E_{i,i} = \sqrt{D_{i,i}}$ for $i = 1, \ldots, n$, so that $D = E^2$. Letting $B = Q^T EQ$, we have $B^2 = Q^T E^2 Q = A$. Furthermore, B is positive semidefinite, since $x^T Bx = (Qx)^T E(Qx) = \sum_{i=1}^n E_{i,i}(Qx)_i^2 \geq 0$ for all $x \in \mathbf{R}^n$.

$4 \Rightarrow 5 \Rightarrow 6$**:** Trivial.

$6 \Rightarrow 1$**:** For any $x \in \mathbf{R}^n$, we have $x^T A x = (Fx)^T(Fx) = \|Fx\|^2 \geq 0$.

$\square$

Similarly, the following holds.

Lemma 6. *Let A be a symmetric real $n \times n$ matrix. The following claims are equivalent.*

- *$A \succ 0$.*
- *A has signature $(n, 0, 0)$.*
- *All eigenvalues of A are positive.*
- *$A = B^2$ for some symmetric positive definite matrix B.*
- *$A = B^2$ for some symmetric regular matrix B.*
- *$A = F^T F$ for some matrix F of rank n.*

Example 1. *Find 5 points $z_1, \ldots, z_5$ in $\mathbf{R}^3$ such that the distance between z_i and z_j is $S_{i,j}$ for the matrix*

$$S = \begin{pmatrix} 0 & 1 & 1 & 1 & \sqrt{3} \\ 1 & 0 & \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 1 & \sqrt{2} & 0 & \sqrt{2} & \sqrt{2} \\ 1 & \sqrt{2} & \sqrt{2} & 0 & \sqrt{2} \\ \sqrt{3} & \sqrt{2} & \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}.$$

Without loss of generality, we can fix $z_1 = (0, 0, 0)$.

We have $S_{ij}^2 = \|z_i - z_j\|^2 = \langle z_i - z_j, z_i - z_j \rangle = \langle z_i, z_i \rangle + \langle z_j, z_j \rangle - 2\langle z_i, z_j \rangle = \|z_i - z_1\|^2 + \|z_j - z_1\|^2 - 2\langle z_i, z_j \rangle = S_{i,1}^2 + S_{j,1}^2 - 2\langle z_i, z_j \rangle$. Let P be the matrix such that $P_{i,j} = \langle z_{i+1}, z_{j+1} \rangle$ for $i, j = 1, \ldots, 4$. Note that $S_{i+1,j+1}^2 = S_{i+1,1}^2 + S_{j+1,1}^2 - 2P_{i,j}$, and thus $P_{i,j} = \frac{1}{2}(S_{i+1,1}^2 + S_{j+1,1}^2 - S_{i+1,j+1}^2)$. Furthermore, $S_{i+1,1}^2 = P_{i,i}$. Hence, the matrix S uniquely determines P, and vice versa. In our example, we have

$$P = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 3 \end{pmatrix}.$$

Let $Z = (z_2|z_3|z_4|z_5)^T$; we have $P = ZZ^T$, and thus $P \succeq 0$. We can express $P = QDQ^T$, where

$$Q = \begin{pmatrix} \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{2} \\ \frac{1}{\sqrt{12}} & 0 & -\frac{2}{\sqrt{6}} & \frac{1}{2} \\ \frac{3}{\sqrt{12}} & 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

is orthogonal and

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

Let

$$E = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

so that $D = EE^T$, and $P = QE(QE)^T$. Hence, we can set $Z = QE$, and thus

$$\begin{aligned} z_2 &= \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right) \\ z_3 &= \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}\right) \\ z_4 &= \left(\frac{1}{\sqrt{3}}, 0, -\frac{2}{\sqrt{6}}\right) \\ z_5 &= \left(\sqrt{3}, 0, 0\right) \end{aligned}$$

is a solution.

Lemma 7. *If $A \succ 0$ and B is regular, then the diagonal entries of BAB^T are positive.*

Proof. We have $(BAB^T)_{i,i} = e_i^T BAB^T e_i = (B^T e_i)^T A (B^T e_i) > 0$, since $B^T e_i \neq o$. $\square$

Suppose we apply the algorithm from Sylvester's law of inertia (simultaneous Gaussian elimination on rows and columns). By Lemma 7, we never need to exchange rows (and columns), and we only need to add multiples of rows (columns) to the latter ones. Also, we only multiply rows (columns) by positive numbers. Consequently, each of the row operations is expressed by a lower-triangular matrix with positive entries on the diagonal. We end up with the identity matrix.

Corollary 8 (Cholesky decomposition)**.** *If A is a positive definite $n \times n$ matrix, then there exists a unique lower-triangular matrix L with positive entries on the diagonal such that $A = LL^T$.*

Proof. The existence follows by the analysis of the simultaneous Gaussian elimination on rows and columns.

For the uniqueness, suppose that $A = LL^T$. Then for any $i \leq j$, we have

$$A_{i,j} = \sum_{k=1}^{n} L_{i,k} L_{j,k} = \sum_{k=1}^{i} L_{i,k} L_{j,k},$$

and thus

$$L_{i,i} L_{j,i} = A_{i,j} - \sum_{k=1}^{i-1} L_{i,k} L_{j,k}.$$

Since the diagonal entries of L are positive, it follows that

$$L_{i,i} = \sqrt{A_{i,i} - \sum_{k=1}^{i-1} L_{i,k}^2}$$

for $i = 1, \ldots, n$, and that

$$L_{j,i} = \frac{A_{i,j} - \sum_{k=1}^{i-1} L_{i,k} L_{j,k}}{L_{i,i}}$$

for $i < j$. Only the entries of L with lexicographically smaller indices appear in these formulas, and thus they uniquely determine L. □

Example 2. *Find the Cholesky decomposition of*

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

Let

$$L = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}.$$

Then

$$LL^T = \begin{pmatrix} a^2 & ab & ad \\ ab & b^2+c^2 & bd+ce \\ ad & bd+ce & d^2+e^2+f^2 \end{pmatrix}.$$

Hence,

$$\begin{aligned} a^2 = 1 &\Rightarrow a = 1 \\ ab = b &= -1 \\ ad = d &= 2 \\ b^2+c^2 = 1+c^2 = 5 &\Rightarrow c = 2 \\ bd+ce = -2+2e = -2 &\Rightarrow e = 0 \\ d^2+e^2+f^2 = 4+f^2 = 5 &\Rightarrow f = 1, \end{aligned}$$

and thus

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Cholesky decomposition is useful to solve systems of linear equations, when the left-hand side happens to be positive definite: To solve $Ax = b$, we first solve $Ly = b$, then $L^T x = y$, by forward and backward substitution.

Example 3. *Let A be the matrix from Example 2. Solve the system of equations $Ax = (3, 5, 7)^T$.*

First, solve the system $Ly = (3, 5, 7)^T$:

$$\begin{aligned} y_1 &= 3 \\ -y_1 + 2y_2 &= 5 \\ 2y_1 + y_3 &= 7 \end{aligned}$$

This gives $y = (3, 4, 1)^T$.

Next, we solve the system $L^T x = y$:

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 3 \\ 2x_2 &= 4 \\ x_3 &= 1 \end{aligned}$$

This gives $x = (3, 2, 1)$.

Corollary 9. *Let A be a symmetric real $n \times n$ matrix. For $i = 1, \ldots, n$, let A_i be the $i \times i$ matrix obtained from A by removing the last $n - i$ rows and columns. The matrix A is positive definite if and only if* $\det(A_i) > 0$ *for* $i = 1, \ldots, n$.

Proof. Perform the simultaneous Gaussian elimination on rows and columns, without ever swapping rows (columns) and with only adding to later rows (columns) and multiplying by positive numbers. This does not affect the sign of $\det(A_i)$ for any i.

If A is positive definite, this is possible and we end up with the identity matrix I, with $\det(I_i) = 1$ for $i = 1, \ldots, n$.

If $\det(A_i) > 0$ for $i = 1, \ldots, n$, this also ensures that the process goes through and ends up with the identity matrix, showing that A has signature $(n, 0, 0)$, and thus $A \succ 0$. □

2 Linear and semidefinite programming

A <u>linear program</u> is a problem of form "minimize $c^T x$ subject to $Ax \geq b$," for a real matrix A and real vectors b and c (or with $\leq/=$ in the constraints, or with max instead of min). Linear programs can be solved exactly in polynomial time.

Example 4. *We have warehouses in towns s_1, ..., s_4, containing* 100*,* 200*,* 250*, and* 100 *kgs of our product, respectively. We have orders from towns t_1, t_2, and t_3, requesting* 150*,* 200*, and* 200 *kgs of our product, respectively. Shipping* 1 *kg of our product from town s_i to town t_j costs $C_{i,j}$ for the following matrix*

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 2 & 2 \\ 4 & 3 & 1 \end{pmatrix}.$$

How to ship the product to minimize the cost?

Let x_{ij} denote the amount of the product sent from the town s_i to the town s_j. Then, we want to minimize

$$\sum_{ij} C_{ij} x_{ij} = x_{11} + 2x_{12} + 3x_{13} + 2x_{21} + x_{22} + 3x_{23} + x_{31} + 2x_{32} + 2x_{33} + 4x_{41} + 3x_{42} + x_{43}$$

subject to

$$\begin{aligned} x_{11}+x_{12}+x_{13} &\le 100 & x_{21}+x_{22}+x_{23} &\le 200\\ x_{31}+x_{32}+x_{33} &\le 250 & x_{41}+x_{42}+x_{43} &\le 100\\ x_{11}+x_{21}+x_{31}+x_{41} &= 150 & x_{12}+x_{22}+x_{32}+x_{42} &= 200\\ x_{13}+x_{23}+x_{33}+x_{43} &= 200 & \textit{for all } i,j,\ x_{ij} &\ge 0 \end{aligned}$$

A semidefinite program is a problem of form "minimize $c^T x$ subject to $A_0 + A_1x_1 + \ldots + A_nx_n \succeq 0$" for real symmetric matrices $A_0, \ldots, A_n$ and a real vector $c \in \mathbf{R}^n$.

Example 5. *The largest eigenvalue of a symmetric real matrix A is the minimum t such that $It - A \succeq 0$.*

We may also require positive semidefiniteness of several matrices (since $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \succeq 0$ if and only if $A \succeq 0$ and $B \succeq 0$), and linear equalities and inequalities among $x_1, \ldots, x_n$ (since $\alpha_1x_1+\ldots+\alpha_nx_n+b \ge 0$ is equivalent to $(\alpha_1)x_1 + \ldots + (\alpha_n)x_n + (b) \succeq 0$). Hence, positive semidefinite programming generalizes linear programming. The solutions to semidefinite programs can be approximated arbitrarily well in polynomial time.

Example 6. *Let G be a graph with vertex set $\{z_1, \ldots, z_n\}$. Let $\theta(G)$ be the minimum r such that there exist vectors $v_1, \ldots, v_n \in \mathbf{R}^n$ such that $\|v_i\| = 1$ for $i = 1, \ldots, n$, and such that $\langle v_i, v_j\rangle \le r$ for all i, j such that $z_iz_j \in E(G)$.*

This minimum can be computed by a semidefinite program. By Lemma 5, an $n \times n$ symmetric matrix X is positive semidefinite if and only if there exists vectors $v_1, \ldots, v_n \in \mathbf{R}^n$ such that $X_{i,j} = \langle v_i, v_j\rangle$ for $1 \le i, j \le n$. Hence, equivalently, we want to minimize r in a semidefinite program with variables r and x_{ij} for $i \le j$, such that $x_{i,i} = 1$ for $i = 1, \ldots, n$, $x_{ij} \le r$ whenever $z_iz_j \in E(G)$, and $X = \sum_{1\le i\le j\le n} A(i,j)x_{ij} \succeq 0$, where $A(i, j)$ is the matrix with $[A(i,j)]_{i,j} = [A(i,j)]_{j,i} = 1$ and all other entries equal to 0.

Let us establish an upper and a lower bound on $\theta(G)$. Firstly, observe that $\theta(K_t) = -\frac{1}{t-1}$ for all $t \ge 1$, and that $\theta(H) \le \theta(G)$ for any subgraph H of G, and thus $\theta(G) \ge -\frac{1}{\omega(G)-1}$. On the other hand, if G can be properly k-colored, then we can assign to all vertices of G of the same color one of the vectors of the optimal solution for K_k, and thus $\theta(G) \le -\frac{1}{\chi(G)-1}$.

Therefore, $\omega(G) \le -\frac{1}{\theta(G)} - 1 \le \chi(G)$, and thus $\theta(G)$ gives an efficiently computable bound for two hard to compute graph parameters.