Applied Linear Algebra

Instructor: Jodin Morey   moreyjc@umn.edu

3.4 Positive Definite Matrices

We've seen a lot of inner products, but is there a general form for them?

Let's find out. Inner product must take in two vectors $\vec{x}, \vec{y}$, and produce a scalar.

Observe, we can always recharacterize $\vec{x}, \vec{y}$ as $\vec{x} = x_1\hat{e}_1 + \ldots + x_n\hat{e}_n = \sum_{i=1}^n x_i\hat{e}_i$, and

$\vec{y} = y_1\hat{e}_1 + \ldots + y_n\hat{e}_n = \sum_{i=1}^n y_i\hat{e}_i$.   (e.g., $(2,1) = 2\hat{e}_1 + 1\hat{e}_2$)

And since every inner product must satisfy the axioms, observe that:

$\langle\vec{x},\vec{y}\rangle = \left\langle \sum_{i=1}^n x_i\hat{e}_i, \sum_{i=1}^n y_i\hat{e}_i \right\rangle = \sum_{i,j=1}^n \langle\hat{e}_i,\hat{e}_j\rangle x_i y_j$.

(e.g., $\langle\vec{x},\vec{y}\rangle = \left\langle x_1\hat{e}_1 + x_2\hat{e}_2,\ y_1\hat{e}_1 + y_2\hat{e}_2 \right\rangle = \langle\hat{e}_1,\hat{e}_1\rangle x_1y_1 + \langle\hat{e}_1,\hat{e}_2\rangle x_1y_2 + \langle\hat{e}_2,\hat{e}_1\rangle x_2y_1 + \langle\hat{e}_2,\hat{e}_2\rangle x_2y_2$)

Therefore, notating $k_{ij} := \langle\hat{e}_i,\hat{e}_j\rangle$, we have: $\langle\vec{x},\vec{y}\rangle = \sum_{i,j=1}^n k_{ij}x_iy_j = \vec{x}^T\mathbf{K}\vec{y}$, where $\mathbf{K}$ is a matrix of the k_{ij}.

Definition: Any inner product can be expressed as a general *bilinear form*: $\vec{x}^T\mathbf{K}\vec{y}$.

Let's use inner product axioms to reveal these properties in bilinear forms.

Due to the symmetry of inner products, we see that $k_{ij} = \langle\hat{e}_i,\hat{e}_j\rangle = \langle\hat{e}_j,\hat{e}_i\rangle = k_{ji}$, therefore $\mathbf{K} = \mathbf{K}^T$ is symmetric.

Using this symmetry, we then find that: $\langle\vec{x},\vec{y}\rangle = \vec{x}^T\mathbf{K}\vec{y} = \left(\vec{x}^T\mathbf{K}\vec{y}\right)^T$ (because a constant is its own transpose)

$= \vec{y}^T\mathbf{K}^T\vec{x} = \vec{y}^T\mathbf{K}\vec{x} = \langle\vec{y},\vec{x}\rangle$.

Regarding positivity, observe: $\left|\vec{x}\right|^2 = \langle\vec{x},\vec{x}\rangle = \vec{x}^T\mathbf{K}\vec{x} = \sum_{i,j=1}^n k_{ij}x_ix_j \geq 0$ for all $\vec{x} \in \mathbb{R}^n$, with equality **iff** $\vec{x} = \vec{0}$.

Definition: $\mathbf{K}^{n\times n}$ is *positive definite* if it is symmetric ($\mathbf{K}^T = \mathbf{K}$), and satisfies *positivity condition*

$\vec{x}^T\mathbf{K}\vec{x} > 0$ for all $\vec{0} \neq \vec{x} \in \mathbb{R}^n$. We sometimes write $\mathbf{K} > 0$ to mean $\mathbf{K}$ is a positive definite.

Therefore…

Theorem: Every inner product on $\mathbb{R}^n$ is given by $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \mathbf{K} \vec{y}$ for $\vec{x}, \vec{y} \in \mathbb{R}^n$ and some positive definite $\mathbf{K}^{n \times n}$.

Definition: Given symmetric $\mathbf{K}$ (not necessarily arising from an inner product), the homogeneous quadratic polynomial $q(\vec{x}) = \sum_{i,j=1}^{n} k_{ij} x_i x_j$ is called a *quadratic form* on $\mathbb{R}^n$. The quadratic form is called *positive definite* if $q(\vec{x}) > 0$ for all $0 \neq \vec{x} \in \mathbb{R}^n$. In other words, q is positive definite **iff** $\mathbf{K}$ is.

Definition: More generally, a quadratic form and its associated symmetric $\mathbf{K}$ are called *positive semi-definite* if $q(\vec{x}) = \vec{x}^T \mathbf{K} \vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^n$ (also written $\mathbf{K} \geq 0$).

Observe that a positive semidefinite matrix implies the possibility of *null directions*.

That is, nonzero vectors $\vec{z}$ such that $q(\vec{z}) = \vec{z}^T \mathbf{K} \vec{z} = 0$.

A positive definite matrix is not allowed null directions, $\ker \mathbf{K} = \left\{ \vec{0} \right\}$, therefore:

Proposition: If a matrix is positive definite, then it is nonsingular.

! Even though $\mathbf{K} = \begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix}$ has two negative entries, it is positive definite.

$$q(\vec{x}) = \vec{x}^T \mathbf{K} \vec{x} = 4x_1^2 - 4x_1 x_2 + 3x_2^2$$

$$= (2x_1 - x_2)^2 + 2x_2^2 \geq 0.$$

! It can be shown from analytic geometry that $ax_1^2 + 2bx_1x_2 + cx_2^2 > 0$ **iff**

$a > 0$ and $ac - b^2 > 0$ (positive leading coefficient & positive determinant).

See evidence for this in book.

Reading the matrix $\mathbf{K}$ directly from a quadratic form.

If given a quadratic form: $ax^2 + bxy + cy^2$, $\mathbf{K}$'s diagonal entries are the coefficients of the associated squared terms.

For mixed terms, split their coefficients in half. These appear in their respective spots in the matrix. For example, the above quadratic form would have $\mathbf{K} = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$. This generalizes to quadratic forms with more variables.

Example: Is $\mathbf{K} = \begin{bmatrix} 5 & 3 \\ 3 & -2 \end{bmatrix}$ positive definite? If so, write down the formula for the associated inner product.

$$|\mathbf{K}| = \begin{vmatrix} 5 & 3 \\ 3 & -2 \end{vmatrix} = -10 - 9 = -19. \qquad \text{So not positive definite.}$$

Example: Is $\mathbf{K} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ positive definite? If so, write down the formula for the associated inner product.

$$|\mathbf{K}| = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = 2, \text{ and } a_{11} = 1 > 0, \text{ so positive definite.}$$

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \mathbf{K} \vec{y} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_1 + 2 x_2 y_2.$$

Gram Matrices

Definition: Let V be an inner product space, and let $\vec{v}_1, \ldots, \vec{v}_n \in V$ (not neccesarily $\mathbb{R}^n$).

The associated *Gram matrix*
$$\mathbf{K} = \begin{bmatrix} \langle \vec{v}_1, \vec{v}_1 \rangle & \langle \vec{v}_1, \vec{v}_2 \rangle & \ldots & \langle \vec{v}_1, \vec{v}_n \rangle \\ \langle \vec{v}_2, \vec{v}_1 \rangle & \langle \vec{v}_2, \vec{v}_2 \rangle & \ldots & \langle \vec{v}_2, \vec{v}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{v}_n, \vec{v}_1 \rangle & \langle \vec{v}_n, \vec{v}_2 \rangle & \ldots & \langle \vec{v}_n, \vec{v}_n \rangle \end{bmatrix} \qquad (**)$$

is the $n \times n$ matrix whose entries are the inner products between the $\vec{v}_i$.

Symmetry of the inner product implies symmetry of the Gram matrix:

$k_{ij} = \langle \vec{v}_i, \vec{v}_j \rangle = \langle \vec{v}_j, \vec{v}_i \rangle = k_{ji}$, and hence $\mathbf{K}^T = \mathbf{K}$.

Theorem: All Gram matrices are positive semi-definite.

The Gram matrix $(**)$ is positive definite **iff** inner product elements $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent.

Proof: To prove positive (semi-) definiteness of $\mathbf{K}$, we need to examine the associated quadratic form:

$$q(\vec{x}) = \vec{x}^T \mathbf{K} \vec{x} = \sum_{i,j=1}^n k_{ij} x_i x_j = \sum_{i,j=1}^n \langle \vec{v}_i, \vec{v}_j \rangle x_i x_j.$$

Bilinearity implies we can assemble this summation into a single inner product:

$$q(\vec{x}) = \left\langle \sum_{i=1}^n x_i \vec{v}_i,\ \sum_{j=1}^n x_j \vec{v}_j \right\rangle \qquad \ldots$$

$= \langle \vec{w}, \vec{w} \rangle = |\vec{w}|^2 \geq 0,$ where $\vec{w} = \sum_{i=1}^{n} x_i \vec{v}_i.$

This proves **K** is positive semidefinite.

Moreover, $q(\vec{x}) = |\vec{w}|^2 > 0$ as long as $\vec{w} \neq \vec{0}$. ....

For positive definite need to show: $\vec{v}_1, \ldots, \vec{v}_n$ are independent **iff** $q(\vec{x}) = 0$ only when $\vec{x} = \vec{0}$.

If $\vec{v}_1, \ldots, \vec{v}_n$ are independent, then $\vec{w} = x_1 \vec{v}_1 + \ldots + x_n \vec{v}_n =$??

$= \vec{0}$ **iff** $x_1 = \ldots = x_n = 0$,

and hence $q(\vec{x}) = 0$ **iff** $\vec{x} = \vec{0}$.

This implies that $q(\vec{x})$ and hence **K** are positive definite. ■

Example: Calculate Gram matrix for $\vec{v}_1 := (1,2,-1)$ and $\vec{v}_2 := (3,0,6)$ using Euclidean inner product. Verify positive definiteness of matrix with its bilinear form.

$$\mathbf{K} = \begin{bmatrix} \vec{v}_1 \cdot \vec{v}_1 & \vec{v}_1 \cdot \vec{v}_2 \\ \vec{v}_2 \cdot \vec{v}_1 & \vec{v}_2 \cdot \vec{v}_2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & -3 \\ -3 & 45 \end{bmatrix}.$$

$q(x_1, x_2) = 6x_1^2 - 6x_1x_2 + 45x_2^2.$

Also, due to $a > 0$ & $ac - b^2 > 0$, so $\mathbf{K} > 0$.

Also, observe $\vec{v}_1, \vec{v}_2$ are linearly independent, so $\mathbf{K} > 0$.

Example: Construct the Gram matrix corresponding to $f(x) = 1$, $g(x) = x$, and $h(x) = x^2$ in $C^0[0,1]$ with the L^2 inner product.

$\langle 1,1 \rangle = |1|^2 = \int_0^1 dx = 1.$ $\quad \langle 1,x \rangle = \int_0^1 x dx = \frac{1}{2}.$ $\quad \langle 1,x^2 \rangle = \int_0^1 x^2 dx = \frac{1}{3}.$ $\quad \langle x,x \rangle = |x|^2 = \int_0^1 x^2 dx = \frac{1}{3}.$

$\langle x,x^2 \rangle = \int_0^1 x^3 dx = \frac{1}{4}.$ $\quad \langle x^2,x^2 \rangle = |x^2|^2 = \int_0^1 x^4 dx = \frac{1}{5}.$

Therefore: $\mathbf{K} = \begin{bmatrix} \langle 1,1\rangle & \langle 1,x\rangle & \langle 1,x^2\rangle \\ \langle 1,x\rangle & \langle x,x\rangle & \langle x,x^2\rangle \\ \langle 1,x^2\rangle & \langle x,x^2\rangle & \langle x^2,x^2\rangle \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}.$

And since the monomial functions $\{1,x,x^2\}$ are linearly independent, $\mathbf{K} > 0$.

One can calculate a Gram matrix using a weighted inner product.

❗ Choice of inner product doesn't change independence of vectors. So new Gram matrix will still be positive definite.

Gram Matrix Decomposition

Note that if we define $\mathbf{A} := \left(\vec{v}_1 \ \ldots \ \vec{v}_n\right)$, then $\mathbf{K} = \mathbf{A}^T\mathbf{A}$.

Example: If $\mathbf{A} := \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 0 \end{bmatrix}$, then with Euclidean inner product,

$$\mathbf{K} = \mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \end{bmatrix}\begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \langle \vec{v}_1, \vec{v}_1\rangle & \langle \vec{v}_1, \vec{v}_2\rangle \\ \langle \vec{v}_2, \vec{v}_1\rangle & \langle \vec{v}_2, \vec{v}_2\rangle \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 3 \\ 3 & 9 \end{bmatrix}, \qquad \text{where } \vec{v}_1 = (1,2,-1) \text{ and } \vec{v}_2 = (3,0,0).$$

Proposition: Given $\mathbf{A}^{m\times n}$, the following are equivalent:

- ♦ $n \times n$ Gram matrix $\mathbf{K} = \mathbf{A}^T\mathbf{A}$ is positive definite.
- ♦ $\mathbf{A}$ has linearly independent columns. (by thm above)
- ♦ $\mathit{rank}\,\mathbf{A} = n \leq m$.
- ♦ $\ker \mathbf{A} = \{0\}$.

How can we use $\mathbf{A}^T\mathbf{A}$ technique on other (non-Euclidean) inner products?

Recall every inner product has form $\langle \vec{v}, \vec{w}\rangle = \vec{v}^T\mathbf{C}\vec{w}$, where $\mathbf{C} > 0$ is symmetric, positive definite.

Therefore, given $\vec{v}_1, \ldots, \vec{v}_n$, the entries of the Gram matrix will be: $k_{ij} = \langle \vec{v}_i, \vec{v}_j\rangle = \vec{v}_i^T\mathbf{C}\vec{v}_j$.

In other words, $\mathbf{K} = \mathbf{A}^T\mathbf{C}\mathbf{A}$, where $\mathbf{A} = \begin{bmatrix} \vec{v}_1 & \ldots & \vec{v}_n \end{bmatrix}$. Therefore:

Theorem: Suppose $\mathbf{A}^{m\times n}$ has linearly independent columns. Suppose $\mathbf{C}^{m\times m}$ is any positive definite matrix.

Then, the Gram matrix $\mathbf{K}^{n\times n} = \mathbf{A}^T\mathbf{C}\mathbf{A}$ is a positive definite matrix.

Proposition: Let $\mathbf{K}^{n\times n} = \mathbf{A}^T\mathbf{C}\mathbf{A}$ be the Gram matrix constructed from $\mathbf{A}^{m\times n}$ and $\mathbf{C}^{m\times m} > 0$.

Then, $\ker\mathbf{K} = \ker\mathbf{A}$, and hence $rank\,\mathbf{K} = rank\,\mathbf{A}$.

Proof: Clearly, if $\mathbf{A}\vec{x} = \vec{0}$, then $\mathbf{K}\vec{x} = \mathbf{A}^T\mathbf{C}\mathbf{A}\vec{x} = \vec{0}$, and so $\ker\mathbf{A} \subset \ker\mathbf{K}$.

Conversely, if $\mathbf{K}\vec{x} = \vec{0}$, then $0 = \vec{x}^T\mathbf{K}\vec{x} = \vec{x}^T\mathbf{A}^T\mathbf{C}\mathbf{A}\vec{x} = \vec{y}^T\mathbf{C}\vec{y}$, where $\vec{y} = \mathbf{A}\vec{x}$.

Since $\mathbf{C} > 0$, this implies $\vec{y} = \vec{0}$, and hence $\vec{x} \in \ker\mathbf{A}$. So, $\ker\mathbf{K} \subset \ker\mathbf{A}$ and $\ker\mathbf{K} = \ker\mathbf{A}$.

Finally, by a previous theorem, $rank\,\mathbf{K} = n - \dim\ker\mathbf{K} = n - \dim\ker\mathbf{A} = rank\,\mathbf{A}$. ■