Quadratic Functions, Optimization, and Quadratic Forms

Robert M. Freund

February, 2004

2004 Massachusetts Institute of Technology.
1 Quadratic Optimization

A quadratic optimization problem is an optimization problem of the form:

\[(QP) : \ \text{minimize} \quad f(x) := \frac{1}{2}x^TQx + c^Tx \]

\[\text{s.t.} \quad x \in \mathbb{R}^n.\]

Problems of the form QP are natural models that arise in a variety of settings. For example, consider the problem of approximately solving an over-determined linear system \(Ax = b\), where \(A\) has more rows than columns. We might want to solve:

\[(P_1) : \ \text{minimize} \quad \|Ax - b\| \]

\[\text{s.t.} \quad x \in \mathbb{R}^n.\]

Now notice that \(\|Ax - b\|^2 = x^TA^T Ax - 2b^TAx + b^Tb\), and so this problem is equivalent to:

\[(P_1) : \ \text{minimize} \quad x^TA^T Ax - 2b^TAx + b^Tb \]

\[\text{s.t.} \quad x \in \mathbb{R}^n,\]

which is in the format of QP.

A symmetric matrix is a square matrix \(Q \in \mathbb{R}^{n \times n}\) with the property that \(Q_{ij} = Q_{ji}\) for all \(i, j = 1, \ldots, n\).
We can alternatively define a matrix $Q$ to be symmetric if

$$Q^T = Q.$$ 

We denote the \textit{identity} matrix (i.e., a matrix with all 1’s on the diagonal and 0’s everywhere else) by $I$, that is,

$$I = \begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{pmatrix},$$

and note that $I$ is a symmetric matrix.

The \textit{gradient} vector of a smooth function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is the vector of first partial derivatives of $f(x)$:

$$\nabla f(x) := \begin{pmatrix}
\frac{\partial f(x)}{\partial x_1} \\
\vdots \\
\frac{\partial f(x)}{\partial x_n}
\end{pmatrix}.$$ 

The \textit{Hessian} matrix of a smooth function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is the matrix of second partial derivatives. Suppose that $f(x) : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, and let

$$[H(x)]_{ij} := \frac{\partial^2 f(x)}{\partial x_i \partial x_j}.$$ 

Then the matrix $H(x)$ is a symmetric matrix, reflecting the fact that

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}.$$ 

A very general optimization problem is:
(GP) : minimize $f(x)$

s.t. $x \in \mathbb{R}^n,$

where $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is a function. We often design algorithms for GP by building a local quadratic model of $f(\cdot)$ at a given point $x = \bar{x}$. We form the gradient $\nabla f(\bar{x})$ (the vector of partial derivatives) and the Hessian $H(\bar{x})$ (the matrix of second partial derivatives), and approximate GP by the following problem which uses the Taylor expansion of $f(x)$ at $x = \bar{x}$ up to the quadratic term.

$$(P_2) : \text{minimize } \tilde{f}(x) := f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T H(\bar{x})(x - \bar{x})$$

s.t. $x \in \mathbb{R}^n.$

This problem is also in the format of QP.

Notice in the general model QP that we can always presume that $Q$ is a symmetric matrix, because:

$$x^T Q x = \frac{1}{2} x^T (Q + Q^T) x$$

and so we could replace $Q$ by the symmetric matrix $\tilde{Q} := \frac{1}{2}(Q + Q^T)$.

Now suppose that

$$f(x) := \frac{1}{2} x^T Q x + c^T x$$

where $Q$ is symmetric. Then it is easy to see that:

$$\nabla f(x) = Q x + c$$

and

$$H(x) = Q.$$
Before we try to solve QP, we first review some very basic properties of symmetric matrices.

2 Convexity, Definiteness of a Symmetric Matrix, and Optimality Conditions

- A function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) is a convex function if
  \[
  f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \text{for all } x, y \in \mathbb{R}^n, \quad \text{for all } \lambda \in [0,1].
  \]

- A function \( f(x) \) as above is called a strictly convex function if the inequality above is strict for all \( x \neq y \) and \( \lambda \in (0,1) \).

- A function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R} \) is a concave function if
  \[
  f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y) \quad \text{for all } x, y \in \mathbb{R}^n, \quad \text{for all } \lambda \in [0,1].
  \]

- A function \( f(x) \) as above is called a strictly concave function if the inequality above is strict for all \( x \neq y \) and \( \lambda \in (0,1) \).

Here are some more definitions:

- \( Q \) is symmetric and positive semidefinite (abbreviated SPSD and denoted by \( Q \succeq 0 \)) if
  \[
  x^T Q x \geq 0 \quad \text{for all } x \in \mathbb{R}^n .
  \]

- \( Q \) is symmetric and positive definite (abbreviated SPD and denoted by \( Q \succ 0 \)) if
  \[
  x^T Q x > 0 \quad \text{for all } x \in \mathbb{R}^n, \quad x \neq 0 .
  \]

**Theorem 1** The function \( f(x) := \frac{1}{2} x^T Q x + c^T x \) is a convex function if and only if \( Q \) is SPSD.
**Proof:** First, suppose that $Q$ is not SPSD. Then there exists $r$ such that $r^T Q r < 0$. Let $x = \theta r$. Then $f(x) = f(\theta r) = \frac{1}{2} \theta^2 r^T Q r + \theta c^T r$ is strictly concave on the subset $\{x \mid x = \theta r\}$, since $r^T Q r < 0$. Thus $f(\cdot)$ is not a convex function.

Next, suppose that $Q$ is SPSD. For all $\lambda \in [0, 1]$, and for all $x, y$,

$$f(\lambda x + (1 - \lambda) y) = f(y + \lambda(x - y))$$

$$= \frac{1}{2}(y + \lambda(x - y))^T Q(y + \lambda(x - y)) + c^T(y + \lambda(x - y))$$

$$= \frac{1}{2} y^T Q y + \lambda(x - y)^T Q y + \frac{1}{2} \lambda^2(x - y)^T Q(x - y) + \lambda c^T x + (1 - \lambda)c^T y$$

$$\leq \frac{1}{2} y^T Q y + \lambda(x - y)^T Q y + \frac{1}{2} \lambda(x - y)^T Q(x - y) + \lambda c^T x + (1 - \lambda)c^T y$$

$$= \frac{1}{2} \lambda x^T Q x + \frac{1}{2} (1 - \lambda) y^T Q y + \lambda c^T x + (1 - \lambda)c^T y$$

$$= \lambda f(x) + (1 - \lambda)f(y) ,$$

thus showing that $f(x)$ is a convex function. □

**Corollary 2** $f(x)$ is strictly convex if and only if $Q \succ 0$.

$f(x)$ is concave if and only if $Q \preceq 0$.

$f(x)$ is strictly concave if and only if $Q \prec 0$.

$f(x)$ is neither convex nor concave if and only if $Q$ is indefinite.

**Theorem 3** Suppose that $Q$ is SPSD. The function $f(x) := \frac{1}{2} x^T Q x + c^T x$ attains its minimum at $x^*$ if and only if $x^*$ solves the equation system:

$$\nabla f(x) = Q x + c = 0 .$$
Proof: Suppose that $x^*$ satisfies $Qx^* + c = 0$. Then for any $x$, we have:

$$f(x) = f(x^* + (x - x^*))$$

$$= \frac{1}{2}(x^* + (x - x^*))^T Q(x^* + (x - x^*)) + c^T(x^* + (x - x^*))$$

$$= \frac{1}{2}(x^*)^T Qx^* + (x - x^*)^T Qx^* + \frac{1}{2}(x - x^*)^T Q(x - x^*) + c^T x^* + c^T (x - x^*)$$

$$= \frac{1}{2}(x^*)^T Qx^* + (x - x^*)^T (Qx^* + c) + \frac{1}{2}(x - x^*)^T Q(x - x^*) + c^T x^*$$

$$= \frac{1}{2}(x^*)^T Qx^* + c^T x^* + \frac{1}{2}(x - x^*)^T Q(x - x^*)$$

$$= f(x^*) + \frac{1}{2}(x - x^*)^T Q(x - x^*)$$

$$\geq f(x^*) ,$$

thus showing that $x^*$ is a minimizer of $f(x)$.

Next, suppose that $x^*$ is a minimizer of $f(x)$, but that $d := Qx^* + c \neq 0$. Then:

$$f(x^* + \alpha d) = \frac{1}{2}(x^* + \alpha d)^T Q(x^* + \alpha d) + c^T(x^* + \alpha d)$$

$$= \frac{1}{2}(x^*)^T Qx^* + \alpha d^T Qx^* + \frac{1}{2}\alpha^2 d^T Qd + c^T x^* + \alpha c^T d$$

$$= f(x^*) + \alpha d^T (Qx^* + c) + \frac{1}{2}\alpha^2 d^T Qd$$

$$= f(x^*) + \alpha d^T d + \frac{1}{2}\alpha^2 d^T Qd .$$

But notice that for $\alpha < 0$ and sufficiently small, that the last expression will be strictly less than $f(x^*)$, and so $f(x^* + \alpha d) < f(x^*)$. This contradicts the supposition that $x^*$ is a minimizer of $f(x)$, and so it must be true that $d = Qx^* + c = 0$.  

Here are some examples of convex quadratic forms:

- $f(x) = x^T x$
• \( f(x) = (x - a)^T(x - a) \)

• \( f(x) = (x - a)^T D(x - a) \), where

\[
D = \begin{pmatrix}
d_1 & & \\
& \ddots & \\
& & d_n
\end{pmatrix}
\]

is a diagonal matrix with \( d_j > 0, j = 1, \ldots, n \).

• \( f(x) = (x - a)^T M^T D M (x - a) \), where \( M \) is a non-singular matrix and \( D \) is as above.

3 Characteristics of Symmetric Matrices

A matrix \( M \) is an orthonormal matrix if \( M^T = M^{-1} \). Note that if \( M \) is orthonormal and \( y = Mx \), then

\[
\|y\|^2 = y^T y = x^T M^T M x = x^T M^{-1} M x = x^T x = \|x\|^2,
\]

and so \( \|y\| = \|x\| \).

A number \( \gamma \in \mathbb{R} \) is an eigenvalue of \( M \) if there exists a vector \( \bar{x} \neq 0 \) such that \( M\bar{x} = \gamma \bar{x} \). \( \bar{x} \) is called an eigenvector of \( M \) (and is called an eigenvector corresponding to \( \gamma \)). Note that \( \gamma \) is an eigenvalue of \( M \) if and only if \((M - \gamma I)\bar{x} = 0, \bar{x} \neq 0 \) or, equivalently, if and only if \( \det(M - \gamma I) = 0 \).

Let \( g(\gamma) = \det(M - \gamma I) \). Then \( g(\gamma) \) is a polynomial of degree \( n \), and so will have \( n \) roots that will solve the equation

\[
g(\gamma) = \det(M - \gamma I) = 0,
\]

including multiplicities. These roots are the eigenvalues of \( M \).

**Proposition 4** If \( Q \) is a real symmetric matrix, all of its eigenvalues are real numbers.
Proof: If \( s = a + bi \) is a complex number, let \( \bar{s} = a - bi \). Then \( \bar{s} \cdot t = \bar{s} \cdot \bar{t} \), \( s \) is real if and only if \( s = \bar{s} \), and \( s \cdot \bar{s} = a^2 + b^2 \). If \( \gamma \) is an eigenvalue of \( Q \), for some \( x \neq 0 \), we have the following chains of equations:

\[
Qx = \gamma x \\
\bar{Q} \cdot \bar{x} = \bar{\gamma} \cdot \bar{x} \\
x^T \bar{Q} \bar{x} = x^T Q \bar{x} = x^T (\bar{\gamma} \bar{x}) = \bar{\gamma} x^T \bar{x}
\]

as well as the following chains of equations:

\[
Qx = \gamma x \\
\bar{x}^T Qx = \bar{x}^T (\gamma x) = \gamma \bar{x}^T \bar{x} \\
x^T \bar{Q} \bar{x} = x^T Q \bar{x} = x^T Qx = \gamma x^T x = \gamma x^T \bar{x}.
\]

Thus \( \bar{\gamma} x^T \bar{x} = \gamma x^T \bar{x} \), and since \( x \neq 0 \) implies \( x^T \bar{x} \neq 0 \), \( \bar{\gamma} = \gamma \), and so \( \gamma \) is real. ■

Proposition 5 If \( Q \) is a real symmetric matrix, its eigenvectors corresponding to different eigenvalues are orthogonal.

Proof: Suppose

\[
Qx_1 = \gamma_1 x_1 \text{ and } Qx_2 = \gamma_2 x_2, \ \gamma_1 \neq \gamma_2.
\]

Then

\[
\gamma_1 x_1^T x_2 = (\gamma_1 x_1)^T x_2 = (Qx_1)^T x_2 = x_1^T Qx_2 = x_1^T (\gamma_2 x_2) = \gamma_2 x_1^T x_2.
\]

Since \( \gamma_1 \neq \gamma_2 \), the above equality implies that \( x_1^T x_2 = 0 \). ■

Proposition 6 If \( Q \) is a symmetric matrix, then \( Q \) has \( n \) (distinct) eigenvectors that form an orthonormal basis for \( \mathbb{R}^n \).

Proof: If all of the eigenvalues of \( Q \) are distinct, then we are done, as the previous proposition provides the proof. If not, we construct eigenvectors
iteratively, as follows. Let \( u_1 \) be a normalized (i.e., re-scaled so that its norm is 1) eigenvector of \( Q \) with corresponding eigenvalue \( \gamma_1 \). Suppose we have \( k \) mutually orthogonal normalized eigenvectors \( u_1, \ldots, u_k \), with corresponding eigenvalues \( \gamma_1, \ldots, \gamma_k \). We will now show how to construct a new eigenvector \( u_{k+1} \) with eigenvalue \( \gamma_{k+1} \), such that \( u_{k+1} \) is orthogonal to each of the vectors \( u_1, \ldots, u_k \).

Let \( U = [u_1, \ldots, u_k] \in \mathbb{R}^{n \times k} \). Then \( QU = [\gamma_1 u_1, \ldots, \gamma_k u_k] \).

Let \( V = [v_{k+1}, \ldots, v_n] \in \mathbb{R}^{n \times (n-k)} \) be a matrix composed of any \( n-k \) mutually orthogonal vectors such that the \( n \) vectors \( u_1, \ldots, u_k, v_{k+1}, \ldots, v_n \) constitute an orthonormal basis for \( \mathbb{R}^n \). Then note that

\[
U^T V = 0
\]

and

\[
V^T Q U = V^T [\gamma_1 u_1, \ldots, \gamma_k u_k] = 0.
\]

Let \( w \) be an eigenvector of \( V^T Q V \in \mathbb{R}^{(n-k) \times (n-k)} \) for some eigenvalue \( \gamma \), so that \( V^T Q V w = \gamma w \), and \( u_{k+1} = V w \) (assume \( w \) is normalized so that \( u_{k+1} \) has norm 1). We now claim the following two statements are true:

(a) \( U^T u_{k+1} = 0 \), so that \( u_{k+1} \) is orthogonal to all of the columns of \( U \), and
(b) \( u_{k+1} \) is an eigenvector of \( Q \), and \( \gamma \) is the corresponding eigenvalue of \( Q \).

Note that if (a) and (b) are true, we can keep adding orthogonal vectors until \( k = n \), completing the proof of the proposition.

To prove (a), simply note that \( U^T u_{k+1} = U^T Vw = 0w = 0 \). To prove (b), let \( d = Q u_{k+1} - \gamma u_{k+1} \). We need to show that \( d = 0 \). Note that \( d = QV w - \gamma V w \), and so \( V^T d = V^T Q V w - \gamma V^T V w = V^T Q V w - \gamma w = 0 \). Therefore, \( d = Ur \) for some \( r \in \mathbb{R}^k \), and so

\[
r = U^T U r = U^T d = U^T Q V w - \gamma U^T V w = 0 - 0 = 0.
\]

Therefore, \( d = 0 \), which completes the proof.

\[\blacksquare\]

**Proposition 7** If \( Q \) is SPSD, the eigenvalues of \( Q \) are nonnegative.
**Proof:** If $\gamma$ is an eigenvalue of $Q$, $Qx = \gamma x$ for some $x \neq 0$. Then $0 \leq x^T Qx = x^T (\gamma x) = \gamma x^T x$, whereby $\gamma \geq 0$. $\blacksquare$

**Proposition 8** If $Q$ is symmetric, then $Q = RDR^T$, where $R$ is an orthonormal matrix, the columns of $R$ are an orthonormal basis of eigenvectors of $Q$, and $D$ is a diagonal matrix of the corresponding eigenvalues of $Q$.

**Proof:** Let $R = [u_1, \ldots, u_n]$, where $u_1, \ldots, u_n$ are the $n$ orthonormal eigenvectors of $Q$, and let

$$D = \begin{pmatrix} \gamma_1 & 0 \\ \vdots & \ddots \\ 0 & \gamma_n \end{pmatrix},$$

where $\gamma_1, \ldots, \gamma_n$ are the corresponding eigenvalues. Then

$$(R^T R)_{ij} = u_i^T u_j = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases},$$

so $R^T R = I$, i.e., $R^T = R^{-1}$.

Note that $\gamma_i R^T u_i = \gamma_i e_i$, $i = 1, \ldots, n$ (here, $e_i$ is the $i$th unit vector). Therefore,

$$R^T Q R = R^T Q[u_1, \ldots, u_n] = R^T [\gamma_1 u_1, \ldots, \gamma_n u_n] = [\gamma_1 e_1, \ldots, \gamma_n e_n] = \begin{pmatrix} \gamma_1 & 0 \\ \vdots & \ddots \\ 0 & \gamma_n \end{pmatrix} = D.$$

Thus $Q = (R^T)^{-1} DR^{-1} = RDR^T$. $\blacksquare$

**Proposition 9** If $Q$ is SPSD, then $Q = M^T M$ for some matrix $M$.

**Proof:** $Q = RDR^T = RD^{1/2} D^{1/2} R^T = M^T M$, where $M = D^{1/2} R$. $\blacksquare$
Proposition 10 If $Q$ is SPSD, then $x^T Q x = 0$ implies $Q x = 0$.

Proof:

$$0 = x^T Q x = x^T M^T M x = (M x)^T (M x) = ||M x||^2 \Rightarrow M x = 0 \Rightarrow Q x = M^T M x = 0.$$

Proposition 11 Suppose $Q$ is symmetric. Then $Q \succeq 0$ and nonsingular if and only if $Q \succ 0$.

Proof:

$(\Rightarrow)$ Suppose $x \neq 0$. Then $x^T Q x \geq 0$. If $x^T Q x = 0$, then $Q x = 0$, which is a contradiction since $Q$ is nonsingular. Thus $x^T Q x > 0$, and so $Q$ is positive definite.

$(\Leftarrow)$ Clearly, if $Q \succ 0$, then $Q \succeq 0$. If $Q$ is singular, then $Q x = 0, x \neq 0$ has a solution, whereby $x^T Q x = 0, x \neq 0$, and so $Q$ is not positive definite, which is a contradiction.

4 Additional Properties of SPD Matrices

Proposition 12 If $Q \succ 0$ ($Q \succeq 0$), then any principal submatrix of $Q$ is positive definite (positive semidefinite).

Proof: Follows directly.

Proposition 13 Suppose $Q$ is symmetric. If $Q \succ 0$ and

$$M = \begin{bmatrix} Q & c \\ c^T & b \end{bmatrix},$$

then $M \succ 0$ if and only if $b > c^T Q^{-1} c$.

Proof: Suppose $b \leq c^T Q^{-1} c$. Let $x = (-c^T Q^{-1}, 1)^T$. Then

$$x^T M x = c^T Q^{-1} c - 2 c^T Q^{-1} c + b \leq 0.$$
Thus $M$ is not positive definite.

Conversely, suppose $b > c^T Q^{-1} c$. Let $x = (y, z)$. Then $x^T M x = y^T Q y + 2 z c^T y + b z^2$. If $x \neq 0$ and $z = 0$, then $x^T M x = y^T Q y > 0$, since $Q \succ 0$. If $z \neq 0$, we can assume without loss of generality that $z = 1$, and so $x^T M x = y^T Q y + 2 c^T y + b$. The value of $y$ that minimizes this form is $y = -Q^{-1} c$, and at this point, $y^T Q y + 2 c^T y + b = -c^T Q^{-1} c + b > 0$, and so $M$ is positive definite.

The $k$th leading principal minor of a matrix $M$ is the determinant of the submatrix of $M$ corresponding to the first $k$ indices of columns and rows.

**Proposition 14** Suppose $Q$ is a symmetric matrix. Then $Q$ is positive definite if and only if all leading principal minors of $Q$ are positive.

**Proof:** If $Q \succ 0$, then any leading principal submatrix of $Q$ is a matrix $M$, where

$$Q = \begin{bmatrix} M & N \\ N^T & P \end{bmatrix},$$

and $M$ must be SPD. Therefore $M = R D R^T = R D R^{-1}$ (where $R$ is orthogonal and $D$ is diagonal), and $\det(M) = \det(D) > 0$.

Conversely, suppose all leading principal minors are positive. If $n = 1$, then $Q \succ 0$. If $n > 1$, by induction, suppose that the statement is true for $k = n - 1$. Then for $k = n$,

$$Q = \begin{bmatrix} M & c \\ c^T & b \end{bmatrix},$$

where $M \in \mathbb{R}^{(n-1) \times (n-1)}$ and $M$ has all its principal minors positive, so $M \succ 0$. Therefore, $M = T^T T$ for some nonsingular $T$. Thus

$$Q = \begin{bmatrix} T^T T & c \\ c^T & b \end{bmatrix}.$$

Let

$$F = \begin{bmatrix} (T^T)^{-1} & 0 \\ -c^T (T^T)^{-1} & 1 \end{bmatrix}.$$
Then
\[
FQF^T = \begin{bmatrix}
(TT)^{-1} -c^T(TT)^{-1}c & 0 \\
-c^T(TT)^{-1}c & 1 \\
\end{bmatrix} \begin{bmatrix}
TT & c \\
c & b \\
\end{bmatrix} \begin{bmatrix}
T^{-1} - (TT)^{-1}c \\
0 & 1 \\
\end{bmatrix}
\]
\[
= \begin{bmatrix}
T & (TT)^{-1}c \\
0 & b - c^T(TT)^{-1}c \\
\end{bmatrix} \begin{bmatrix}
T^{-1} - (TT)^{-1}c \\
0 & 1 \\
\end{bmatrix} = \begin{bmatrix}
I & 0 \\
0 & b - c^T(TT)^{-1}c \\
\end{bmatrix}.
\]
Then \( \det Q = \frac{b - c^T(TT)^{-1}c}{\det(F)^2} > 0 \) implies \( b - c^T(TT)^{-1}c > 0 \), and so \( Q \succ 0 \) from Proposition 13.

5 Quadratic Forms Exercises

1. Suppose that \( M \succ 0 \). Show that \( M^{-1} \) exists and that \( M^{-1} \succ 0 \).

2. Suppose that \( M \succeq 0 \). Show that there exists a matrix \( N \) satisfying \( N \succeq 0 \) and \( N^2 := NN = M \). Such a matrix \( N \) is called a “square root” of \( M \) and is written as \( M^{\frac{1}{2}} \).

3. Let \( \|v\| \) denote the usual Euclidian norm of a vector, namely \( \|v\| := \sqrt{v^Tv} \). The operator norm of a matrix \( M \) is defined as follows:
\[
\|M\| := \max_x \{\|Mx\| \mid \|x\| = 1\}.
\]

Prove the following two propositions:

**Proposition 1:** If \( M \) is \( n \times n \) and symmetric, then
\[
\|M\| = \max_{\lambda} \{\|\lambda\| \mid \lambda \text{ is an eigenvalue of } M\}.
\]

**Proposition 2:** If \( M \) is \( m \times n \) with \( m < n \) and \( M \) has rank \( m \), then
\[
\|M\| = \sqrt{\lambda_{\text{max}}(MM^T)},
\]
where \( \lambda_{\text{max}}(A) \) denotes the largest eigenvalue of a matrix \( A \).
4. Let \( \|v\| \) denote the usual Euclidian norm of a vector, namely \( \|v\| := \sqrt{v^T v} \). The operator norm of a matrix \( M \) is defined as follows:

\[
\|M\| := \max_{x \neq 0} \left\{ \|Mx\| \mid \|x\| = 1 \right\}.
\]

Prove the following proposition:

**Proposition:** Suppose that \( M \) is a symmetric matrix. Then the following are equivalent:

(a) \( h > 0 \) satisfies \( \|M^{-1}\| \leq \frac{1}{h} \)
(b) \( h > 0 \) satisfies \( \|Mv\| \geq h \cdot \|v\| \) for any vector \( v \)
(c) \( h > 0 \) satisfies \( |\lambda_i(M)| \geq h \) for every eigenvalue \( \lambda_i(M) \) of \( M \), \( i = 1, \ldots, m \).

5. Let \( Q \succeq 0 \) and let \( S := \{ x \mid x^T Q x \leq 1 \} \). Prove that \( S \) is a closed convex set.

6. Let \( Q \succeq 0 \) and let \( S := \{ x \mid x^T Q x \leq 1 \} \). Let \( \gamma_i \) be a nonzero eigenvalue of \( Q \) and let \( u^i \) be a corresponding eigenvector normalized so that \( \|u^i\|_2 = 1 \). Let \( a^i := \frac{\gamma_i}{\sqrt{\gamma}} \). Prove that \( a^i \in S \) and \( -a^i \in S \).

7. Let \( Q > 0 \) and consider the problem:

\[(P) : \quad z^* = \max_x \quad c^T x \]
\[
\text{s.t.} \quad x^T Q x \leq 1 .
\]

Prove that the unique optimal solution of \((P)\) is:

\[
x^* = \frac{Q^{-1} c}{\sqrt{c^T Q^{-1} c}}
\]

with optimal objective function value

\[
z^* = \sqrt{c^T Q^{-1} c}.
\]
8. Let $Q \succ 0$ and consider the problem:

$$(P): \quad z^* = \max_{x} \; c^T x \quad \text{s.t.} \quad x^T Q x \leq 1.$$ 

For what values of $c$ will it be true that the optimal solution of $(P)$ will be equal to $c$? (Hint: think eigenvectors.)

9. Let $Q \succeq 0$ and let $S := \{x \mid x^T Q x \leq 1\}$. Let the eigendecomposition of $Q$ be $Q = RDR^T$ where $R$ is orthonormal and $D$ is diagonal with diagonal entries $\gamma_1, \ldots, \gamma_n$. Prove that $x \in S$ if and only if $x = Rv$ for some vector $v$ satisfying

$$\sum_{j=1}^{n} \gamma_j v_i^2 \leq 1.$$ 

10. Prove the following:

**Diagonal Dominance Theorem:** Suppose that $M$ is symmetric and that for each $i = 1, \ldots, n$, we have:

$$M_{ii} \geq \sum_{j \neq i} |M_{ij}|.$$ 

Then $M$ is positive semidefinite. Furthermore, if the inequalities above are all strict, then $M$ is positive definite.

11. A function $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is a **norm** if:

   (i) $f(x) \geq 0$ for any $x$, and $f(x) = 0$ if and only if $x = 0$

   (ii) $f(\alpha x) = |\alpha| f(x)$ for any $x$ and any $\alpha \in \mathbb{R}$, and

   (iii) $f(x + y) \leq f(x) + f(y)$.

Define $f_Q(x) = \sqrt{x^T Q x}$. Prove that $f_Q(x)$ is a norm if and only if $Q$ is positive definite.

12. If $Q$ is positive semi-definite, under what conditions (on $Q$ and $c$) will $f(x) = \frac{1}{2} x^T Q x + c^T x$ attain its minimum over all $x \in \mathbb{R}^n$?, be unbounded over all $x \in \mathbb{R}^n$?
13. Consider the problem to minimize $f(x) = \frac{1}{2}x^TQx + c^Tx$ subject to $Ax = b$. When will this program have an optimal solution?, when not?

14. Prove that if $Q$ is symmetric and all its eigenvalues are nonnegative, then $Q$ is positive semi-definite.

15. Let $Q = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$. Note that $\gamma_1 = 1$ and $\gamma_2 = 2$ are the eigenvalues of $Q$, but that $x^TQx < 0$ for $x = (2, -3)^T$. Why does this not contradict the result of the previous exercise?

16. A quadratic form of the type $g(y) = \sum_{j=1}^p y_j^2 + \sum_{j=p+1}^n d_j y_j + d_{n+1}$ is a separable hybrid of a quadratic and linear form, as $g(y)$ is quadratic in the first $p$ components of $y$ and linear (and separable) in the remaining $n - p$ components. Show that if $f(x) = \frac{1}{2}x^TQx + c^Tx$ where $Q$ is positive semi-definite, then there is an invertible linear transformation $y = T(x) = Fx + g$ such that $f(x) = g(y)$ and $g(y)$ is a separable hybrid, i.e., there is an index $p$, a nonsingular matrix $F$, a vector $g$ and constants $d_p, \ldots, d_{n+1}$ such that

$$g(y) = \sum_{j=1}^p (Fx + g)_j^2 + \sum_{j=p+1}^n d_j(Fx + g)_j + d_{n+1} = f(x).$$

17. An $n \times n$ matrix $P$ is called a projection matrix if $P^T = P$ and $PP = P$. Prove that if $P$ is a projection matrix, then

- a. $I - P$ is a projection matrix.
- b. $P$ is positive semidefinite.
- c. $\|Px\| \leq \|x\|$ for any $x$, where $\|\|$ is the Euclidean norm.

18. Let us denote the largest eigenvalue of a symmetric matrix $M$ by “$\lambda_{\max}(M)$.” Consider the program

$$(Q) : \quad z^* = \max_x \quad x^T M x$$

s.t. \quad $\|x\| = 1$, where $M$ is a symmetric matrix. Prove that $z^* = \lambda_{\max}(M)$. 
19. Let us denote the smallest eigenvalue of a symmetric matrix $M$ by "$\lambda_{\text{min}}(M)$." Consider the program

$$(P): \quad z_* = \text{minimum}_x \quad x^T M x$$

s.t. \quad \|x\| = 1 ,

where $M$ is a symmetric matrix. Prove that $z_* = \lambda_{\text{min}}(M)$.

20. Consider the matrix

$$M = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} ,$$

where $A, B$ are symmetric matrices and $A$ is nonsingular. Prove that $M$ is positive semi-definite if and only if $C - B^T A^{-1} B$ is positive semi-definite.