Sec 7.3 Constrained optimization

It's often useful to find the maximum or minimum value of a quadratic form $Q(\vec{x})$ over a set of inputs $\vec{x}$.

__Ex 1__ (When Q has no cross-product terms, it's easy to find the maximum and minimum of $Q(\vec{x})$ subject to the constraint $\|\vec{x}\| = 1$)

Let $Q(\vec{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$.

Find the maximum and minimum values of $Q(\vec{x})$ for $\|\vec{x}\| = 1$.

__Note__ $\|\vec{x}\| = 1$ iff $\vec{x}^T\vec{x} = x_1^2 + x_2^2 + x_3^2 = 1$

__Sol__ * If $\|\vec{x}\| = 1$,

$$\begin{aligned} Q(\vec{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

so 9 is an upper bound of $Q(\vec{x})$ for $\|\vec{x}\| = 1$

Also $Q\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = 9(1^2) + 4(0^2) + 3(0^2) = 9$,

so 9 is the __maximum__ value of $Q(\vec{x})$ for $\|\vec{x}\| = 1$

* Similarly, $Q(\vec{x}) \geqslant 3$ for $\|\vec{x}\| = 1$

So 3 is a lower bound of $Q(\vec{x})$ for $\|\vec{x}\| = 1$

Also $Q\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = 9(0^2) + 4(0^2) + 3(1^2) = 3$

so 3 is the __minimum__ value of $Q(\vec{x})$ for $\|\vec{x}\| = 1$.

__Observe__ • The matrix of the quadratic form Q is

$A = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, and $A\vec{e}_1 = 9\vec{e}_1$, $A\vec{e}_2 = 4\vec{e}_2$, $A\vec{e}_3 = 3\vec{e}_3$

So 9, 4, 3 are eigenvalues of A

• The greatest eigenvalue = the constrained maximum of $Q(\vec{x})$

• – " – least – " – = – " – minimum of $Q(\vec{x})$

Ex 2 Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, and let $Q(\vec{x}) = \vec{x}^T A \vec{x}$ for $\vec{x}$ in $\mathbb{R}^2$

- Graph of $Q(\vec{x})$:
- The points are $\left(x_1, x_2, Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\right)$ for all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

$\downarrow$

- Graph of the constrained values of $Q(\vec{x})$, i.e. the intersection of $z = 3x_1^2 + 7x_2^2$ and $1 = x_1^2 + x_2^2$
- The points are $\left(x_1, x_2, Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)\right)$ for unit vectors $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

$\downarrow$

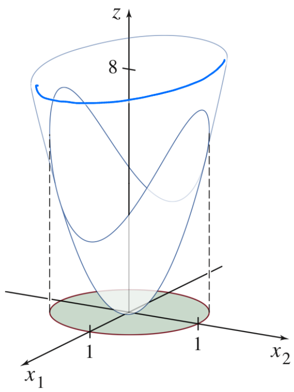


FIGURE 1 $z = 3x_1^2 + 7x_2^2$.

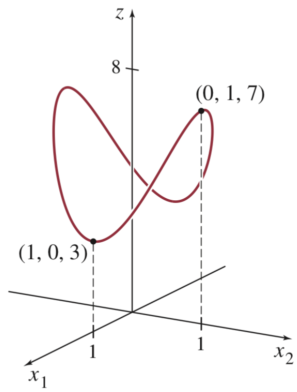


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

See Desmos Link

* The highest points on the intersection curve are $(0,1,7)$ and $(0,-1,7)$.
 — " — lowest — " — $(1,0,3)$ and $(-1,0,3)$.

These corresponds to eigenvalue 7 of A w/ eigenvectors $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$

and eigenvalue 3 of A w/ eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

* Every point on the intersection curve has a z-coordinate between 3 and 7.

* For every number t between 3 and 7, there is a unit vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = t$.

* I.e. $\{$all possible values of $\vec{x}^T A \vec{x}$ for $\|\vec{x}\| = 1\} = [3, 7]$

Fact The set of all possible values of a quadratic form $Q(\vec{x}) = \vec{x}^T A \vec{x}$ for $\|\vec{x}\| = 1$ is equal to a closed interval $[m, M]$

$$m \overset{\text{notation}}{=} \min\{\vec{x}^T A \vec{x} : \|\vec{x}\| = 1\}, \quad M \overset{\text{notation}}{=} \max\{\vec{x}^T A \vec{x} : \|\vec{x}\| = 1\}$$

> **Thm 6** Let A be a symmetric matrix, let $Q(\vec{x}) = \vec{x}^T A \vec{x}$.
>
> Let $m \overset{\text{notation}}{=} \min\{\vec{x}^T A \vec{x} : \|\vec{x}\| = 1\}, \quad M \overset{\text{notation}}{=} \max\{\vec{x}^T A \vec{x} : \|\vec{x}\| = 1\}$.
>
> 1) Then M is the greatest eigenvalue of A and m is the least eigenvalue of A.
>
> 2) $Q(\vec{x}) = \vec{x}^T A \vec{x} = M$ when $\vec{x}$ is a unit M-eigenvector
> $Q(\vec{x}) = \vec{x}^T A \vec{x} = m$ when $\vec{x}$ is a unit m-eigenvector

Proof Orthogonally diagonalize $A = PDP^{-1} = PDP^T$

Do the change of variable $\vec{x} = P\vec{y}$ so that

$$\vec{x}^T A \vec{x} = \underbrace{(P\vec{y})^T A (P\vec{y}) = \vec{y}^T (P^T A P)\vec{y}}_{\text{See Sec 7.2}} = \vec{y}^T D \vec{y}$$

Note $\|\vec{x}\| = \|P\vec{y}\| = \|\vec{y}\|$ for all $\vec{y}$ because orthogonal matrices preserve length (Thm 7 Sec 6.2)

In particular, $\|\vec{y}\| = 1$ iff $\|\vec{x}\| = 1$.

So $\{\vec{x}^T A \vec{x} : \|\vec{x}\| = 1\} = \{\vec{y}^T D \vec{y} : \|\vec{y}\| = 1\}$

(For simplicity, assume 3x3 instead of $n \times n$)

Suppose A is a 3x3 matrix w/ eigenvalues $a \geq b \geq c$ w/ unit eigenvectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$

let $P = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}, \quad D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$

Then
$$\begin{aligned} \vec{y}^T D \vec{y} &= a\, y_1^2 + b\, y_2^2 + c\, y_3^2 \\ &\leq a\, y_1^2 + a\, y_2^2 + a\, y_3^2 \quad \text{since } a \geq b \geq c \text{ and } y_2^2, y_3^2 \geq 0 \\ &= a\|\vec{y}\| = a \end{aligned}$$

Thus $M \leq a$

Moreover, $\vec{y}^T D \vec{y} = a$ when $\vec{y} = \begin{bmatrix}1\\0\\0\end{bmatrix}$, so $M = a$, proving (1)

Note $P \overset{\vec{y}②}{\begin{bmatrix}1\\0\\0\end{bmatrix}} = \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \vec{u}_3\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix} = \overset{\vec{x}②}{\vec{u}_1}$

So $Q(\vec{u}_1) = \vec{u}_1^{\,T} A \vec{u}_1 = [1\,0\,0]\, D \begin{bmatrix}1\\0\\0\end{bmatrix} = a = M$

So $\vec{x}$ is an M-eigenvector of A, proving (2).

Similarly, $m = c$ (smallest eigenvalue of A)

and $m = \vec{x}^T A \vec{x}$ when $\vec{x} = P\begin{bmatrix}0\\0\\1\end{bmatrix} = u_3$ $\square$

<u>Ex 3</u> Let $A = \begin{bmatrix}3&2&1\\2&3&1\\1&1&4\end{bmatrix}$ w/ characteristic equation $0 = -(\lambda-6)(\lambda-3)(\lambda-1)$

Find : (1) the maximum value M of the quadratic form $\vec{x}^T A \vec{x}$ subject to the constraint $\vec{x}^T\vec{x} = 1$

(2) a unit vector at which this maximum value is attained.

<u>Sol</u> (1) The largest eigenvalue of A is $\lambda = 6$, so $\boxed{M = 6.}$

(2) This constrained maximum of $\vec{x}^T A \vec{x}$ is attained when $\vec{x}$ is a unit 6-eigenvector $\vec{u}$ of A.

To find $\vec{u}$, solve for $(A - 6\,Id)\vec{x} = \vec{0}$

$$\left[\begin{array}{ccc|c}3-6&2&1&0\\2&3-6&1&0\\1&1&4-6&0\end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c}-3&2&1&0\\2&-3&1&0\\1&1&-2&0\end{array}\right] \to \left[\begin{array}{ccc|c}1&0&-1&0\\0&1&-1&0\\0&0&0&0\end{array}\right] \quad \begin{array}{l}x_1 - x_3 = 0\\ x_2 - x_3 = 0\end{array}$$

So $\begin{bmatrix}1\\1\\1\end{bmatrix}$ is a 6-eigenvector. $\boxed{\text{Set } \vec{u} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} \text{ or } -\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}.}$

Check: $Q(\vec{u}) = 3\left(\frac{1}{\sqrt{3}}\right)^2 + 3\left(\frac{1}{\sqrt{3}}\right)^2 + 4\left(\frac{1}{\sqrt{3}}\right)^2 + 2(2)\left(\frac{1}{\sqrt{3}}\right)^2 + 2(1)\left(\frac{1}{\sqrt{3}}\right)^2 + 2(1)\left(\frac{1}{\sqrt{3}}\right)^2$

$= 18\,\frac{1}{3} = 6$ ✓

Thm 7 Let A be a symmetric matrix.

Let λ_1 be a largest eigenvalue of A, and $\vec{u}_1$ be a unit eigenvector corresponding to λ_1.

(1) Then the maximum value of $\vec{x}^T A \vec{x}$ subject to the constraints $\underbrace{\vec{x}^T \vec{x} = 1}_{\text{unit vector}}$ and $\underbrace{\vec{x}^T \vec{u}_1 = 0}_{\vec{x} \text{ is orthogonal to } \vec{u}_1}$ is the second greatest eigenvalue, λ_2.

(2) This maximum is attained when $\vec{x}$ is a unit eigenvector $\vec{u}_2$ corresponding to λ_2.

Example 4 (Idea of proof for Thm 7 when A is a diagonal matrix)

Let $Q(\vec{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$. (The quadratic form from Example 1)

Find the maximum value of $Q(\vec{x})$ subject to the constraints

(Extra)

$$\vec{x}^T \vec{x} = 1 \quad \text{and} \quad \vec{x}^T \vec{u}_1 = 0$$

where $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, a unit 9-eigenvector (recall 9 is the greatest eigenvalue of the matrix of the quadratic form).

Sol The constraint $[x_1\, x_2\, x_3] \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 0$ means $x_1 = 0$

$$\begin{aligned} \text{So} \quad Q(\vec{x}) &= 9(0^2) + 4(x_2^2) + 3(x_3^2) \\ &\leq 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4 \end{aligned}$$

for $\vec{x}$ w/ these two constraints

We have $Q\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = 4$ so 4 is a maximum.

Note 4 is the second greatest eigenvalue and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is a 4-eigenvector.

EXAMPLE 5 Let A be the matrix in Example 3 and let $\mathbf{u}_1$ be a unit eigenvector corresponding to the greatest eigenvalue of A. Find the maximum value of $\mathbf{x}^TA\mathbf{x}$ subject to the conditions

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0 \tag{4}$$

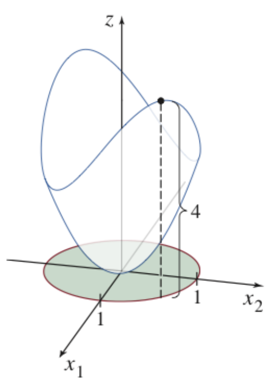
SOLUTION From Example 3, the second greatest eigenvalue of A is $\lambda = 3$. Solve $(A - 3I)\mathbf{x} = \mathbf{0}$ to find an eigenvector, and normalize it to obtain

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

The vector $\mathbf{u}_2$ is automatically orthogonal to $\mathbf{u}_1$ because the vectors correspond to different eigenvalues. Thus the maximum of $\mathbf{x}^TA\mathbf{x}$ subject to the constraints in (4) is 3, attained when $\mathbf{x} = \mathbf{u}_2$. ■

Practice Problems

1. Let $Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Find a change of variable that transforms Q into a quadratic form with no cross-product term, and give the new quadratic form.
2. With Q as in Problem 1, find the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which the maximum is attained.



The maximum value of $Q(\mathbf{x})$ subject to $\mathbf{x}^T\mathbf{x} = 1$ is 4.

Solutions to Practice Problems

1. The matrix of the quadratic form is $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. It is easy to find the eigenvalues, 4 and 2, and corresponding unit eigenvectors, $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. So the desired change of variable is $\mathbf{x} = P\mathbf{y}$, where $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (A common error here is to forget to normalize the eigenvectors.) The new quadratic form is $\mathbf{y}^TD\mathbf{y} = 4y_1^2 + 2y_2^2$.
2. The maximum of $Q(\mathbf{x})$, for a unit vector $\mathbf{x}$, is 4 and the maximum is attained at the unit eigenvector $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. [A common incorrect answer is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. This vector maximizes the quadratic form $\mathbf{y}^TD\mathbf{y}$ instead of $Q(\mathbf{x})$.]

— end of PDF —