Some topics in orthogonality

MATH 133 - Vectors, Matrices and Geometry

1. Orthogonal complements and orthogonal projections

Lemma 1.0.1. *Let $W \subseteq \mathbb{R}^n$ be a subspace. We say that a vector $v \in \mathbb{R}^n$ is orthogonal to W if $v \cdot w = 0$ for all $w \in W$. Let $W^\perp = \{v \in \mathbb{R}^n : v \cdot w = 0 \text{ for all } w \in W\}$ then $W^\perp$ is a subspace of $\mathbb{R}^n$. If $W = \mathrm{Span}\{w_1, \dots, w_k\}$ then $v \in W^\perp$ if and only if $v \cdot w_i = 0$ for all $i = 1, \dots k$.*

(proved in class)

Theorem 1.0.2. *Let $W = \mathrm{Span}\{w_1, \dots, w_k\} \subseteq \mathbb{R}^n$ be a subspace. Then $W^\perp$ is the solutions to the homogenous system defined by the matrix $A = \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \\ \vdots \\ \overline{w_k} \end{pmatrix}$. Thus,*

$$\dim(W^\perp) = n - \dim(W).$$

(proved in class)

Theorem 1.0.3. *Let $W \subseteq \mathbb{R}^n$ be a subspace. Every vector $v \in \mathbb{R}^n$ can be written uniquely as $v = v_1 + v_2$ with $v_1 \in W$ and $v_2 \in W^\perp$. Define the projection on W to be the map*

$$\mathrm{proj}_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad \mathrm{proj}_W(v) = v_1.$$

Then proj_W is a linear map with the following properties:

1. *if $v \in W$ then $\mathrm{proj}_W(v) = v$;*
2. *if $v \in W^\perp$ then $\mathrm{proj}_W(v) = 0$;*
3. *if $\{w_1, \dots, w_k\}$ is an orthogonal basis for W then for any v we have*

$$\mathrm{proj}_W(v) = \frac{v \cdot w_1}{\|w_1\|^2} w_1 + \frac{v \cdot w_2}{\|w_2\|^2} w_2 + \cdots + \frac{v \cdot w_k}{\|w_k\|^2} w_k.$$

(proved in class)

Theorem 1.0.4. *We have $(W^\perp)^\perp = W$.*

Proof. Let $w \in W$ and $u \in W^\perp$ then $w \cdot u = 0$. This holds for every $u \in W^\perp$ thus $w \in (W^\perp)^\perp$. We conclude that $W \subseteq (W^\perp)^\perp$. However, we also know that $\dim(W) + \dim(W^\perp) = n$ and $\dim(W^\perp) + (W^\perp)^\perp = n$. It follows that $\dim(W) = \dim((W^\perp)^\perp)$ and hence $W = (W^\perp)^\perp$. $\square$

Application: If $W = \mathrm{Span}\{w_1, \ldots, w_k\}$ we can describe W as the solutions to a system of linear equations as follows. We calculate a basis $u_1, \ldots, u_r$ to $W^\perp$ by solving the system $\begin{pmatrix} \underline{w_1} \\ \underline{w_2} \\ \vdots \\ \underline{w_k} \end{pmatrix}$. Then, by the same reasoning W (viewed as $(W^\perp)^\perp$) is the solutions to the system of linear equations $\begin{pmatrix} \underline{u_1} \\ \underline{u_2} \\ \vdots \\ \underline{u_k} \end{pmatrix}$. (One has to be very careful about row and columns here).

Example: suppose that $W = \mathrm{Span}\{(1,2,1,-1),(1,0,1,3),(3,2,3,5)\}$. Then we have to solve the system defined by $\begin{pmatrix} 1 & 2 & 1 & -1 \\ 1 & 0 & 1 & 3 \\ 3 & 2 & 3 & 5 \end{pmatrix}$. This matrix has REF $\begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. A basis to the solutions is given by $u_1 = (-1,0,1,0), u_2 = (-3,2,0,1)$. It follows that W is exactly the solutions of $\begin{pmatrix} -1 & 0 & 1 & 0 \\ -3 & 2 & 0 & 1 \end{pmatrix}$. That is, W is defined by the equations $-x+z=0, -3x+2y+w=0$.

Example: For which conditions b_1, b_2, b_3 can we solve the system

$$3x_1 + x_2 + x_3 = b_1, \quad 2x_1 + x_3 = b_2, \quad 11x_1 + 5x_2 + 3x_3 = b_3.$$

The question is therefore: when is the column vector $(b_1, b_2, b_3)^T$ in the span of the columns of the matrix $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 0 & 1 \\ 11 & 5 & 3 \end{pmatrix}$. We write the columns as rows; we wish to recognize then the row space of $\begin{pmatrix} 3 & 2 & 11 \\ 1 & 0 & 5 \\ 1 & 1 & 3 \end{pmatrix}$. Row reduction gives the matrix $\begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$. Therefore the column space, call it W is spanned by $\left\{ \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \right\}$. We calculate that $W^\perp$ is the span of $(-5,2,1)$. Therefore, we can solve the equations if and only if $-5b_1 + 2b_2 + b_3 = 0$.

2. The Gram-Schmidt process

Theorem 2.0.5. *Let $x_1, \ldots, x_k$ be a basis for a subspace W of $\mathbb{R}^n$. Let $W_1 = \mathrm{Span}(x_1), W_2 = \mathrm{Span}(x_1, x_2), \ldots, W_k = \mathrm{Span}(x_1, x_2, \ldots, x_k)$. Define the following vectors:*

$$\begin{aligned} v_1 &= x_1 \\ v_2 &= x_2 - \tfrac{x_2 \cdot v_1}{\|v_1\|^2} \cdot v_1 \\ v_3 &= x_3 - \tfrac{x_3 \cdot v_1}{\|v_1\|^2} \cdot v_1 - \tfrac{x_3 \cdot v_2}{\|v_2\|^2} \cdot v_2 \\ &\vdots \\ v_k &= x_k - \tfrac{x_k \cdot v_1}{\|v_1\|^2} \cdot v_1 - \tfrac{x_k \cdot v_2}{\|v_2\|^2} \cdot v_2 - \cdots - \tfrac{x_k \cdot v_{k-1}}{\|v_{k-1}\|^2} \cdot v_{k-1} \end{aligned}$$

Then, for each $i = 1, \ldots, k$ we have that $v_1, v_2, \ldots, v_i$ is an orthogonal basis for W_i. In particular, $\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \ldots, \frac{v_k}{\|v_k\|}\}$ is an orthonormal basis for W.

Example: Find the orthogonal projection of $(1,1,1)$ of $W = \{(x,y,z) : x + 2y - 5z = 0\}$.

We find first any basis for W by finding a basis for the solution space of a homogenous system. We get that $u_1 = (-2,1,0), u_2 = (5,0,1)$ is a basis for W. Apply Gram-Schmidt:

$v_1 = u_1 = (-2,1,0)$,

$v_2 = u_2 - \frac{u_2 \cdot v_1}{\|v_1\|^2} \cdot v_1 = (5,0,1) + \frac{10}{5}(-2,1,0) = (1,2,1)$.

Then v_1, v_2 is an orthogonal basis and we may apply our formula for the projection:

$\mathrm{proj}_W((1,1,1)) = \frac{(1,1,1)\cdot v_1}{\|v_1\|^2} \cdot v_1 + \frac{(1,1,1)\cdot v_2}{\|v_2\|^2} \cdot v_2 = \frac{-1}{5}(-2,1,0) + \frac{4}{6}(1,2,1) = \frac{1}{15}(16,17,10)$.

Example: Find an orthogonal basis for $\mathbb{R}^3$ containing the vector $(1,1,1)$.

We first just complete it to a basis, getting for example $(1,1,1),(0,1,0),(0,0,1)$. (Compute the determinant, which is 1, to see this is a basis.) Now apply Gram-Schmidt:

$v_1 = (1,1,1)$

$v_2 = (0,1,0) - \frac{(0,1,0)\cdot(1,1,1)}{\|(1,1,1)\|^2}(1,1,1) = (0,1,0) - \frac{1}{3}(1,1,1) = \frac{1}{3}(-1,2,-1)$.

$v_3 = (0,0,1) - \frac{(0,0,1)\cdot(1,1,1)}{\|(1,1,1)\|^2}(1,1,1) - \frac{(0,0,1)\cdot\frac{1}{3}(-1,2,-1)}{\|\frac{1}{3}(-1,2,-1)\|^2}(-1,2,-1)$

$= (0,0,1) - \frac{1}{3}(1,1,1) + \frac{1}{6}(-1,2,-1) = \frac{1}{2}(-1,0,1)$.

3. Orthogonal diagonalization and symmetric matrices

We say that an $n \times n$ matrix A with real coefficients is orthogonally diagonalizable (OD) if there is an orthogonal matrix Q such that

$$Q^T\ A\ Q = D$$

is a diagonal matrix. Note that since Q is orthogonal, this is the same as saying that $Q^{-1}AQ = D$ is diagonal. Thus, this is a more sophisticated version of the problem of diagonalization.

Proposition 3.0.6. *If A is OD then A is symmetric.*

Proof. Indeed, we have that $Q^T\ A\ Q = D$. Take transpose to find that $(Q^T\ A\ Q)^T = D^T = D$. Thus, $Q^T\ A\ Q = (Q^T\ A\ Q)^T = Q^T A^T (Q^T)^T = Q^T A^T Q$. Since Q is invertible, we get that $A = A^T$. $\square$

The following lemma is used in the proof of the theorem following it.

Lemma 3.0.7. *If A is a symmetric real matrix then every eigenvalue of A is real.*

Proof. Suppose that λ is a (perhaps complex) eigenvalue and $v = (z_1, \dots, z_n)^T$ a non-zero eigenvector of λ. Then, on the one hand, $\bar{v}^T(Av) = \bar{v}(\lambda v) = \lambda \bar{v}^T v$. On the other hand, $\bar{v}^T(Av) = \bar{v}^T A^T v = (A\bar{v})^T v = (\overline{Av})^T v = (\overline{\lambda v})^T v = \bar{\lambda}\bar{v}^T v$. Note that $\bar{v}^T v = \sum_{i=1}^n |z_i|^2 \neq 0$. It follows that $\lambda = \bar{\lambda}$ and so λ is real. $\square$

Theorem 3.0.8. *(The Spectral Theorem) Let A be a symmetric real matrix, then A is orthogonally diagonalizable.*

Before discussing the algorithm for OD-ing a matrix, we prove the following Proposition (that explains why the algorithm works).

Proposition 3.0.9. *Let A be a symmetric real matrix and let λ_1, λ_2 be distinct eigenvalues of A. Then*

$$E_{\lambda_1} \perp E_{\lambda_2}.$$

(That is, every vector in the first space is orthogonal to every vector in the second space).

Proof. Let $v_i \in E_{\lambda_i}$ be a non-zero eigenvector. Then $\lambda_1 v_1 \cdot v_2 = \lambda_1 v_1^T v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T (Av_2) = v_1^T \lambda_2 v_2 = \lambda_2 v_1 \cdot v_2$. Since $\lambda_1 \neq \lambda_2$, it follows that $v_1 \cdot v_2 = 0$. $\square$

Orthogonal diagonalization of any real symmetric matrix

Goal: Given an $n \times n$ real symmetric matrix A, to find an orthogonal real matrix Q so that $Q^T AQ = D$ is diagonal.

(1) Calculate the characteristic polynomial $f(x) = \det(A - xI_n)$ of A.

(2) Write $f(x) = (-1)^n (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2} \cdots (x-\lambda_r)^{m_r}$. Note that $m_1+m_2+\cdots+m_r = n$; the λ_i are the eigenvalues of A and are real numbers.

(3) Calculate the eigenspace $E_{\lambda_i} = \{v : (A - \lambda_i I_n)v = 0\}$ for every eigenvalue λ_i; for every eigenvalue λ_i we have $\dim(E_{\lambda_i}) = m_i$.

(4) Calculate a basis $\mathscr{C}_i$ for every eigenspace E_{λ_i}. Using Gram-Schmidt calculate from $\mathscr{C}_i$ an orthonormal basis $\mathscr{B}_i$ for E_{λ_i}. Let $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 \cup \cdots \cup \mathscr{B}_r$. Then each $\mathscr{B}_i$ consists of m_i vectors and $\mathscr{B} = \{v_1, \ldots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

(5) Let $Q = (v_1|v_2|\cdots|v_n)$. Then Q is an orthogonal matrix and $Q^T AQ$ is the diagonal matrix with blocks

$$\begin{pmatrix} m_1\boxed{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}} & & & \\ & m_2\boxed{\begin{matrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{matrix}} & & \\ & & \ddots & \\ & & & m_r\boxed{\begin{matrix}\lambda_r & & \\ & \ddots & \\ & & \lambda_r\end{matrix}} \end{pmatrix}$$

Note that we know that A is diagonalizable. Thus the algorithm does not need to determine whether A is diagonalizable or not as it had to do in the case of a general matrix. Also, the last proposition guarantees that Q is an orthogonal matrix.

Example: OD the matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 2 & -2 \end{smallmatrix}\right)$.

The characteristic polynomial of A is $x^2 + x - 6 = (x-2)(x-3)$. We find that $E_2 = \mathrm{Span}\{(2,1)^T\}, E_{-3} = \mathrm{Span}\{(1,-2)^T\}$. Thus, $\mathscr{B}_1 = \{\frac{1}{\sqrt{5}}(2,1)^T\}, \mathscr{B}_2 = \{\frac{1}{\sqrt{5}}(1,-2)^T\}$ and $\mathscr{B} = \{\frac{1}{\sqrt{5}}(2,1)^T, \frac{1}{\sqrt{5}}(1,-2)^T\}$. Note that this is indeed an orthonormal basis. The matrix Q is just $\frac{1}{\sqrt{5}}\left(\begin{smallmatrix} 2 & 1 \\ 1 & -2 \end{smallmatrix}\right)$.

Example: OD the matrix $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

The characteristic polynomial of A is $-(x-1)(x+1)(x-2)$. The eignespaces are $E_1 = \mathrm{Span}\{(1,0,-1)^T\}, E_{-1} = \mathrm{Span}\{(-1,2,-1)^T\}, E_2 = \mathrm{Span}\{(1,1,1)^T\}$. We get an orthonormal basis $\mathscr{B} = \{\frac{1}{\sqrt{2}}(1,0,-1)^T, \frac{1}{\sqrt{6}}(-1,2,-1)^T, \frac{1}{\sqrt{3}}(1,1,1)^T\}$. The matrix Q is $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$.

Example: OD the matrix $A = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix}$.

The characteristic polynomial is $(x-1)^3(x-3)$. The eigenspace E_1 is the solutions of $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ and a basis for it is $\{(0,1,0,0)^T, (0,0,1,0)^T, (1,0,0,-1)^T\}$. Note that this is an orthogonal basis and normalize it to get $\mathscr{B}_1 = \{(0,1,0,0)^T, (0,0,1,0)^T, \frac{1}{\sqrt{2}}(1,0,0,-1)^T\}$. The eigenspace E_3 is $\mathrm{Span}\{\frac{1}{\sqrt{2}}(1,0,0,1)^T\}$ and $\mathscr{B}_2 = \{\frac{1}{\sqrt{2}}(1,0,0,1)^T\}$. An orthonormal basis of eigenvectors if given by $\mathscr{B} = \{(0,1,0,0)^T, (0,0,1,0)^T, \frac{1}{\sqrt{2}}(1,0,0,-1), \frac{1}{\sqrt{2}}(1,0,0,1)^T\}$.

The matrix Q is thus $\begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$.

4. Quadratic forms

A quadratic form in the variables $x_1, \ldots, x_n$ is a function f of the form

$$f(x_1, \ldots, x_n) = \sum_{i \le j} c_{ij} x_i x_j.$$

We may write such a function as

$$f(x_1,\dots,x_n)=(x_1,\dots,x_n)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

where the matrix A has diagonal elements $c_{11},\dots,c_{nn}$, is symmetric, and its (i,j) element is $c_{ij}/2$.

Example: $x_1^2+2x_1x_2+5x_2^5=(x_1,x_2)\left(\begin{smallmatrix}1&1\\1&5\end{smallmatrix}\right)\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)$. Similarly, $2x_1^2+x_2x_3+x_3^2-4x_1x_3$ is $(x_1,x_2,x_3)\begin{pmatrix}2&0&-2\\0&0&1/2\\-2&1/2&1\end{pmatrix}\left(\begin{smallmatrix}x_1\\x_2\\x_3\end{smallmatrix}\right)$.

We would like to change the variables so that the expression simplifies. For example, we would like to know if this expression is always positive or not and similar questions.

Note that if we define new set of variable $Q\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$. Then, in the new sets of variables our expression is

$$f(y_1,\dots,y_n)=(y_1,y_2,\dots,y_n)Q^TAQ\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

Let us then choose Q orthogonal such that $Q^TAQ=D=\mathrm{diag}[d_1,\dots,d_n]$ is a diagonal matrix. Then we conclude that with the set of variables $y_1,\dots,y_n$ the quadratic form has the shape

$$f(y_1,\dots,y_n)=(y_1,y_2,\dots,y_n)D\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2.$$

Note that the d_i are the eigenvalues of A. We thus conclude the following theorem:

Theorem 4.0.10. *(The Principal Axes Theorem) Let* $(x_1,\dots,x_n)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ *be a quadratic form.*

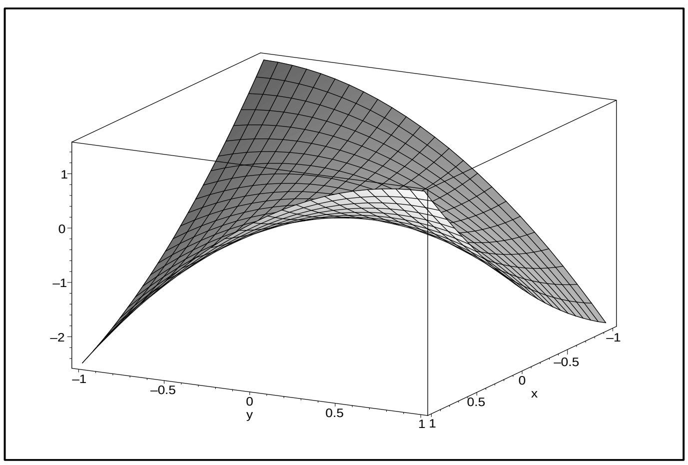
(1) *There is an orthogonal matrix Q such that with respect to the variables* $Q\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ *the quadratic form is written as $d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2$, the d_i being the eigenvalues of A.*

(2) *The quadratic form is positive-definite (resp. negative definite, resp. indefinite), that is, has always positive values (resp. negative values, resp. positive and negative values) if and only if each eigenvalue d_i is positive (resp. negative, resp. some are positive and some are negative).*

(3) *The maximum (resp. minimum) of the quadratic form over vectors x with $\|x\| = 1$ is the maximal (resp. minimal) eigenvalue.*

One application is the following. Let $z = f(x, y)$ be a differentiable function in two variables. Suppose that at some point (x_0, y_0) the derivatives $\partial f/\partial x, \partial f/\partial y$ vanish. Consider the symmetric matrix $\begin{pmatrix} \partial^2 f/\partial x^2 & \partial^2 f/\partial x\partial y \\ \partial^2 f/\partial x\partial y & \partial^2 f/\partial y^2 \end{pmatrix}$ at the point (x_0, y_0). One proves in calculus of several variables that If this matrix is negative definite then (x_0, y_0) is a maximum; if it is positive definite then (x_0, y_0) is a minimum; if it is indefinite then (x_0, y_0) is a saddle point. Note that by the Principal Axes Theorem this can be checked by simply checking the eigenvalues.

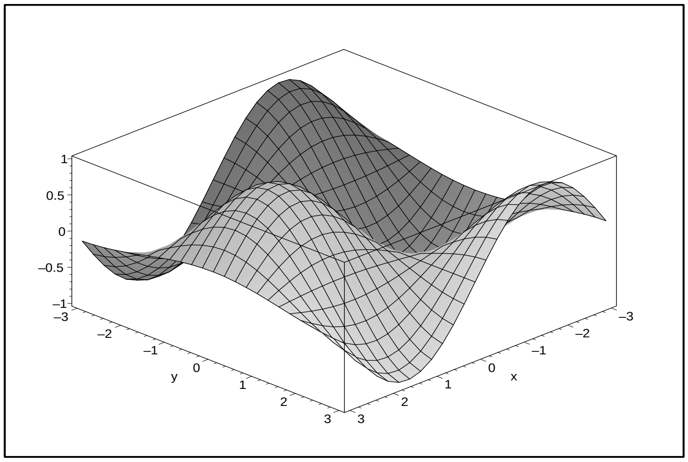
Example: Consider the function $z = x^2/2 + 2xy - y^2$. (Here it is best to ignore the fact that this is a quadratic form and simply think of it as an example of a function of two variables.) Determine the nature of the point $(0, 0)$. The partials are $x + 2y, 2x - 2y$, which vanish at $(0, 0)$. The matrix of second derivatives is $A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$. The eigenvalues of it are $2, -3$. It follows that $(0, 0)$ is a saddle point.



Similarly, consider the function $z = sin(x)cos(y)$, whose partials are $cos(x)cos(y), -sin(x)sin(y)$. Those vanish at the point $x = \pi/2, y = 0$. The matrix of second derivatives at the point is $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. This is negative definite and we conclude that $(\pi/2, 0)$ is a maximum.

Example: For the quadratic form $x^2 - 6xy + y^2$, find a change of variables such that the form has the shape $d_1 X^2 + d_2 Y^2$. What are d_1, d_2? What is the maximum of the form subject to the condition $x^2 + y^2 = 1$?

The matrix representing the form is $A = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}$. The characteristic polynomial is $x^2 - 2x - 8 = (x - 4)(x + 2)$. We find that the eigenspaces are spanned by $\frac{1}{\sqrt{2}}(1, -1)$ and



$\frac{1}{\sqrt{2}}(1,1)$. The matrix Q is thus $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$. If we let $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = Q \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}}(X+Y) \\ \frac{1}{\sqrt{2}}(-X+Y) \end{pmatrix}$ (so $\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right) = Q^{-1} \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = Q^T \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \begin{pmatrix} \frac{1}{\sqrt{2}}(x-y) \\ \frac{1}{\sqrt{2}}(x+y) \end{pmatrix}$) and substitute that into the equation we shall get $4X^2 - 2Y^2$. Thus the maximum is 4 and the minimum is -2.