# Risks, Classification, and Pattern Recognition[1]

Ningsong Shen

*Western University, London, Ontario N6A 5B7, Canada*

November 27, 2020

**Abstract.** A working file for applications in quantifying convexity.

*Keywords and phrases:* Hessian, matrices, eigenvalues, convexity

# 1 Introduction

Convexity plays an important role in many areas of research and practice. For example, consider an application of convexity with the automated harvesting of apples (Kelman and Linker, 2014). The use of convexity there to detect apples does not suffer adverse effects from poor lighting or varying colours. However, a different challenge is posed, as apples are not perfectly convex. Another area of use is in financial mathematics, where convexity is often used to measure risk. A simple example would be looking at bond risk with curves. In fact, there are many objects in the world and many applications where the data is not perfectly convex but where it would be useful in knowing the degree to which it is not. The following text will examine a method to quantify the convexity of a function or object from Davydov, Moldavskaya, and Zitikis (2019) and apply it to different fields. The beginning sections will look at the indices of (non)convexity for a matrix. The middle sections will examine the application of the method on convex and non-convex objects. Finally, the conclusion will discuss the different future applications and potential avenues that this new method will open up.

# 2 Finding the Eigenvalues of a Matrix

We start with an analysis of fundamental matrix techniques, which are critical in the method for quantifying convexity.

[1]This research has been supervised by Ričardas Zitikis and supported by the Natural Sciences and Engineering Research Council Undergraduate Student Research Award (NSERC USRA).

**Definition 2.1.** $\lambda$ is called an eigenvalue of square matrix $A$ if there is a nonzero vector $x$ such that $Ax = \lambda x$.

This can be rewritten as $(A - \lambda I)x = 0$ which helps find the eigenvalues for a given vector $x$. We recall a property of singular matrices, which have a determinant of 0. Therefore, $\lambda$ is an eigenvalue of $A$ if and only if $A - \lambda I$ is singular, in other words, if $\det(A - \lambda I) = 0$.

**Definition 2.2.** The determinant of a 2 x 2 matrix with entries $x_{ij}$ for $i, j = 1, 2$ is $x_{11}x_{22} - x_{12}x_{21}$

We can use the definition of a determinant to calculate the determinant for $A - \lambda I$ in the $2 \times 2$ case with

$$(x_{11} - \lambda)(x_{22} - \lambda) - x_{12}x_{21} = 0,$$

which can be rearranged as follows

$$\lambda^2 - (x_{11} + x_{22})\lambda + x_{11}x_{22} - x_{12}x_{21} = 0.$$

This is of the form $ax^2 + bx + c = 0$ which can be easily solved with the quadratic formula, and hence we present the following theorem.

**Theorem 2.1.** *The eigenvalues of a $2 \times 2$ matrix with entries $x_{ij}$ for $i, j = 1, 2$ can be denoted as follows using the entries of the matrix:*

$$\lambda = \frac{(x_{11} + x_{22}) \pm \sqrt{(x_{11} + x_{22})^2 - 4(x_{11}x_{22} - x_{12}x_{21})}}{2}$$

Using this general form, we can calculate eigenvalues for a continuous function in multidimensional spaces.

# 3 Convexity in the $2 \times 2$ Case

To illustrate, we select the differentiable function $f(x, y) = \sin(xy)$ plotted in Figure 3.1.

The $2 \times 2$ Hessian of $f(x, y) = \sin(xy)$ was computed and its positive and negative eigenvalues were plotted in Figure 3.2. This is enough information to plot the normalized index of lack of convexity (Figure 3.3) and index of convexity (Figure 3.4).

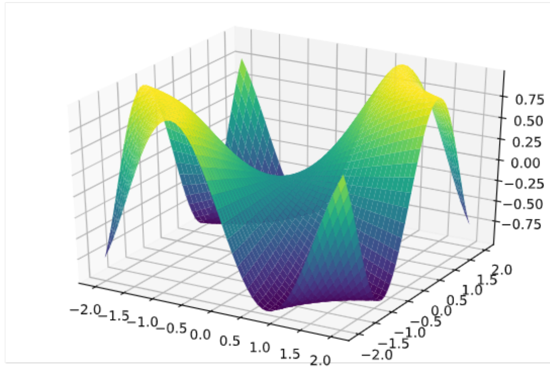
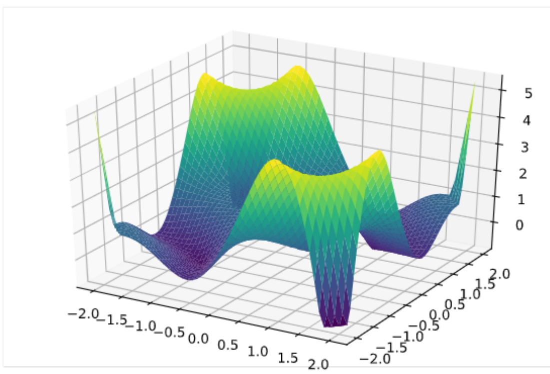
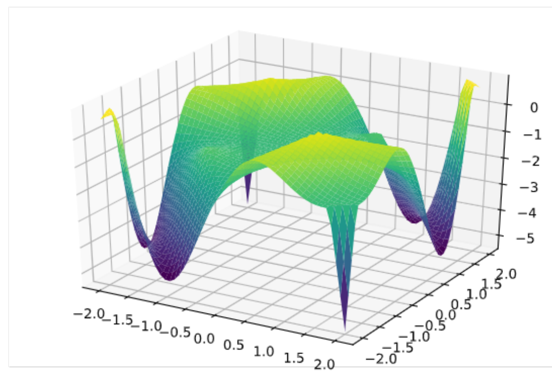



Figure 3.1: Function $f(x, y) = \sin(xy)$




(a) Positive




(b) Negative

Figure 3.2: Eigenvalues of the $2 \times 2$ Hessian, $\lambda_1$ and $\lambda_2$

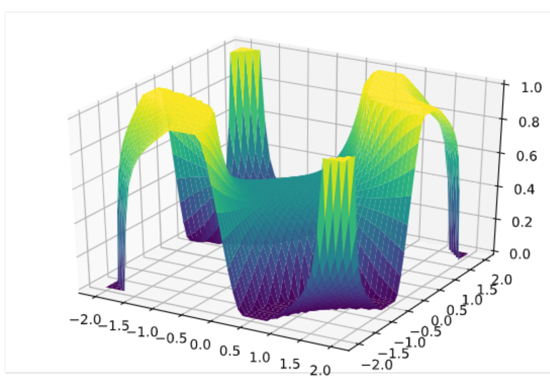



Figure 3.3: The normalized index of lack of convexity $\text{NLOC}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^-(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

Other functions were also plotted. Of note are the functions that would not graph properly, likely due to entries of the Hessian being 0. This would cause issues when

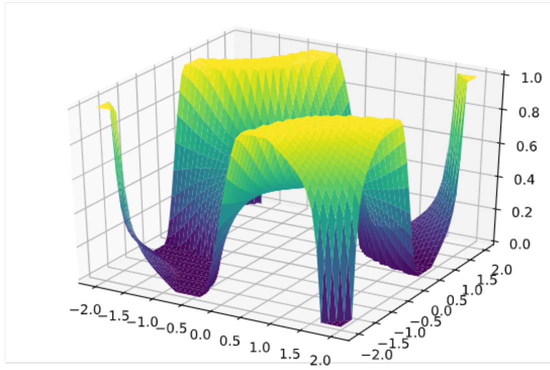



Figure 3.4: The index of convexity $\mathrm{CONV}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^+(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

graphing and calculating the functions of 2 variables. Another plotting solution may be required, but sticking to functions with non-disappearing partial derivatives (e.g. $e^{xy}$) will work well.

# 4 The Hessian, its Eigenvalues, and Convexity

Now we move to the Hessian, which is often thought of as a slope for multivariate functions. Since convexity is related to slope, it is important to note some details about the Hessian.

**Definition 4.1.** For a square $n \times n$ matrix, there will be a maximum of $n$ eigenvalues.

Therefore, for a $2 \times 2$ Hessian matrix, there will never be more than 2 eigenvalues. We define a few more terms for the upcoming discussion on the significance of the Hessian matrix.

**Definition 4.2.** A convex function is a function between two points $a$ and $b$ which are joined by a line segment that will fall on or above the function. Note that a linear function is considered convex.

**Definition 4.3.** A strictly convex function is a convex function which has one local minimum.

**Definition 4.4.** $a$ is a positive number if $a \geq 0$ while $a$ is a negative number if $a \leq 0$.

**Definition 4.5.** $b$ is a strictly positive number if $b > 0$ while $b$ is a strictly negative number if $b < 0$. Note the exclusion of 0 in both cases.

The Hessian matrix is a matrix of second partial derivatives. Usually, mixed partial derivatives are equivalent. Thus, the Hessian matrix be regarded as symmetric, and such matrices always have real eigenvalues. This can be proved by rewriting $x_{12} = x_{21}$ and using the discriminant from Theorem 2.1.

$$\begin{aligned}\Delta &= (x_{11} + x_{22})^2 - 4(x_{11}x_{22} - x_{12}^2) \\ &= x_{11}^2 + 2x_{11}x_{22} + x_{22}^2 - 4x_{11}x_{22} + 4x_{12}^2 \\ &= (x_{11} - x_{22})^2 + 4x_{12}^2\end{aligned}$$

All the terms are positive, hence the discriminant is positive, and the eigenvalues must be real. If the discriminant is strictly positive, there will be two distinct eigenvalues, otherwise, there will only be one eigenvalue with algebraic multiplicity 2. Figure 4.1 plots locations of only one eigenvalue for $f(x, y) = \sin(xy)$.

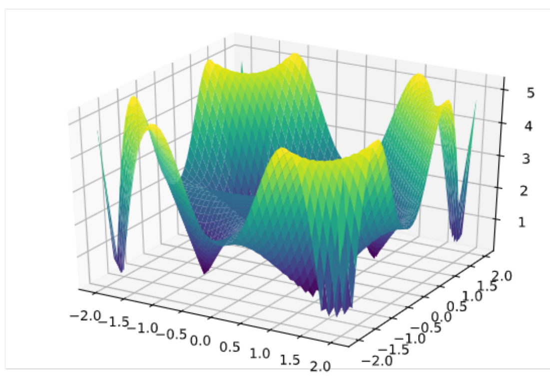



Figure 4.1: Locations where there only one eigenvalue exists

For a single-variable function, the second derivative indicates curvature. Regions where the second derivative are strictly positive indicate strict convexity. In the $2 \times 2$ case, the Hessian matrix is used as the equivalent of the second derivative representing curvature and convexity. The ideas of being strictly positive need to be extended to matrices.

**Definition 4.6.** A symmetric $n \times n$ matrix $A$ is positive-definite if $z^T Az$ is strictly positive for every non-zero column vector $z$. The same matrix is positive semi-definite if $z^T Az$ is positive for every non-zero column vector $z$.

Regions of a multivariable function where the Hessian matrix is positive-definite must then be strictly convex, while regions where the Hessian matrix is positive semi-definite must be convex. Other regions will have a lack of convexity.

## 5 Canonical Decomposition

The canonical decomposition is need to show relationships between terms in the method of Davydov, Moldavskaya, and Zitikis (2019).

**Example 5.1.** Here we will look at the decomposition of a matrix. We start with the following symmetric $2 \times 2$ matrix:

$$H = \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}.$$

This matrix will be diagonalized orthogonally. Using Theorem 2.1, we can deduce that the eigenvalues are $\pm 2$ and that the respective eigenvectors are

$$\begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ 1 \end{bmatrix}$$

These vectors are orthogonal to each other, but when put together as columns of a matrix, they do not form an orthogonal matrix because they are not normalized. Using the Gram-Schmidt process, we obtain a normalized matrix:

$$Q = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

This matrix is orthogonal and its inverse is equal to its transpose.

Now we look at the diagonal matrix of eigenvalues

$$\Lambda = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix},$$

which means that

$$\Lambda^+ = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$

and

$$\Lambda^- = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}.$$

Using these matrices, we can calculate $H^+$ and $H^-$.

$$H^+ = Q\Lambda^+Q^T = \begin{bmatrix} \frac{3}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$$

$$H^- = Q\Lambda^-Q^T = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{3}{2} \end{bmatrix}$$

From earlier, we can obtain that $H = Q\Lambda Q^T$ is true. Similarly, $H^+ = Q\Lambda^+Q^T$ holds, $H^- = Q\Lambda^+Q^T$ too, and so

$$H = H^+ - H^- = \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}.$$

The nuclear norm of $H$ is

$$\|H\|_* = \sum_{i=1}^{d} |\lambda_i| = 4,$$

similarly $\|H^+\|_* = 2$, $\|H^-\|_* = 2$, and so $\|H\|_* = \|H^+\|_* + \|H^-\|_*$ holds. Thus, we are now ready to confidently use the method for quantifying convexity.

# 6 Half a Unit Ball

As a preamble to the main application of the indices to apples and lemons, we will investigate the case of convexity for a simpler object: the 2-D unit ball.

The upper half of a unit ball centered at $(0, 0)$ was drawn on blank paper. Definition from Wikipedia. Proof modified from note.

**Definition 6.1.** Let S be a vector space or an affine space over the real numbers, or more generally, over some ordered field. This includes Euclidean spaces, which are affine spaces. A subset $C$ of $S$ is convex if, for all $x$ and $y$ in $C$, the line segment connecting $x$ and $y$ is included in $C$. This means that the affine combination $(1 - t)x + ty$ belongs to $C$, for all $x$ and $y$ in C, and $t$ in the interval $[0, 1]$.

Let $C = (x, y) \in R^2 : x^2 + y^2 \leq 1, y \geq 0$ be the upper half of the unit ball. The set $C$ is convex if for all $a, b \in C$, the points $ta + (1 - t)b \in C$ for all $t \in [0, 1]$. Since $a, b \in C$, we have

$$a_1^2 + a_2^2 \leq 1, a_2 \geq 0$$

$$ta_1^2 + ta_2^2 \le t, a_2 \ge 0,$$

$$b_1^2 + b_2^2 \le 1, b_2 \ge 0,$$

$$(1-t)b_1^2 + (1-t)b_2^2 \le (1-t), b_2 \ge 0.$$

Adding the equations together,

$$ta_1^2 + (1-t)b_1^2 + ta_2^2 + (1-t)b_2^2 \le t + (1-t), a_2 \ge 0, b_2 \ge 0,$$

we use the fact that $f(x) = x^2$ is convex:

$$(ta_1 + (1-t)b_1)^2 + (ta_2 + (1-t)b_2)^2 \le 1, a_2 \ge 0, b_2 \ge 0$$

proving that the upper half of a unit ball is convex.

We also show that the boundary of the upper half of a unit ball is concave using the definition of concavity.

**Definition 6.2.** A real-valued function $f$ on an interval (or more generally, a convex set in vector space) is said to be concave if, for any $x$ and $y$ in the interval and any $t \in [0,1]$, $f((1-t)x + ty) \ge (1-t)f(x) + tf(y)$.

Let $f(x) = \sqrt{1-x^2}$ be the equation for the boundary of the upper half of the unit ball. $f$ is concave if

$$\sqrt{1 - ((1-t)x + ty)^2} \ge (1-t)\sqrt{1-x^2} + t\sqrt{1-y^2}$$

and since both sides are positive,

$$\begin{aligned}
1 - ((1-t)x + ty)^2 &\ge (1-t)^2(1-x^2) + 2t(1-t)\sqrt{(1-x^2)(1-y^2)} + t^2(1-y^2) \\
1 - (1-t)^2x^2 - 2t(1-t)xy - t^2y^2 &\ge (1-t)^2 - (1-t)^2x^2 + 2t(1-t)\sqrt{(1-x^2)(1-y^2)} + t^2 - t^2y^2 \\
1 - 2t(1-t)xy &\ge (1-t)^2 + 2t(1-t)\sqrt{(1-x^2)(1-y^2)} + t^2 \\
1 &\ge 1 - 2t + 2t^2 + 2t(1-t)\sqrt{(1-x^2)(1-y^2)} + 2t(1-t)xy \\
0 &\ge -2t + 2t^2 + 2t(1-t)(\sqrt{(1-x^2)(1-y^2)} + xy) \\
0 &\ge -1 + t + (1-t)(\sqrt{(1-x^2)(1-y^2)} + xy) \\
0 &\ge (t-1)(1 - \sqrt{(1-x^2)(1-y^2)} + xy)
\end{aligned}$$

which is true because $(t-1)$ is always negative and $(1 - \sqrt{(1-x^2)(1-y^2)} + xy)$ is always positive. This proves that $f$ is concave.

# 7 Half a 3D Unit Ball

We will also look at the 3-D unit ball, which is an even closer relative to the apple or lemon. This was particularly helpful in reasoning through the apple and lemon proofs.

The upper half of a unit ball centered at $(0, 0, 0)$ was drawn on blank paper. The convexity of this ball will be proved in a similar manner to the above. It is of note (in both cases) that the set itself is convex, however the surface containing the set is concave.

Let $C = (x, y, z) \in R^2 : x^2 + y^2 + z^2 \leq 1, z \geq 0$ be the upper half of the 3D unit ball. The set $C$ is convex if for all $a, b \in C$, the points $ta + (1-t)b \in C$ for all $t \in [0, 1]$. Since $a, b \in C$, we have

$$a_1^2 + a_2^2 + a_3^2 \leq 1, a_3 \geq 0$$

$$ta_1^2 + ta_2^2 + ta_3^2 \leq t, a_3 \geq 0,$$

$$b_1^2 + b_2^2 + b_3^2 \leq 1, b_3 \geq 0,$$

$$(1-t)b_1^2 + (1-t)b_2^2 + (1-t)b_3^2 \leq (1-t), b_3 \geq 0.$$

Adding the equations together,

$$ta_1^2 + (1-t)b_1^2 + ta_2^2 + (1-t)b_2^2 + ta_3^2 + (1-t)b_3^2 \leq t + (1-t), a_3 \geq 0, b_3 \geq 0.$$

we use the fact that $f(x) = x^2$ is convex:

$$(ta_1 + (1-t)b_1)^2 + (ta_2 + (1-t)b_2)^2 + (ta_3 + (1-t)b_3)^2 \leq 1, a_3 \geq 0, b_3 \geq 0.$$

proving that the upper half of a 3D unit ball is convex.

We also show that the boundary of the upper half of a 3D unit ball is concave using through the composition of concave functions. Suppose $f$ is a concave and non-decreasing function, and $g$ is a concave function. We want to show that $f(g(x))$ is concave:

$$f(g(tx + (1-t)y)) \geq tf(g(x)) + (1-t)f(g(y)).$$

Since $f$ is non decreasing and $g$ is concave,

$$f(g(tx + (1-t)y)) \geq f(tg(x) + (1-t)g(y))$$

and since $f$ is a concave function:

$$f(tg(x)+(1-t)g(y)) \geq tf(g(x)+(1-t)f(g(y))$$

proving that the composition of $f(g(x))$ is concave. We have the following theorem:

**Theorem 7.1.** *Let $h(x) = f(g(x))$. The composite function $h$ is concave if $f$ is non-decreasing and concave, and $g$ is concave.*

Let $f(\mathbf{x}) = \sqrt{1-x_1^2-x_2^2}$ be the equation for the boundary of the upper half of the 3D unit ball. We can decompose $f$ as $g(h(x))$ like this:

$$g(x) = \sqrt{x}$$

which is concave and non-decreasing, and

$$h(\mathbf{x}) = 1-x_1^2-x_2^2$$

which is concave. Using Theorem 7.1, the resulting function is also concave. This proves that $f$ is concave.

Using these two examples of unit balls, we can apply similar principles to proofs involving more complex objects like the mathematical apple or lemon (which will be described below).

# 8 Apples and Lemons

In many cases, it is quite useful to classify convex objects from non-convex objects and especially to determine the areas of convexity compared to those of non-convexity. Modern artificial intelligence techniques often rely on convex restrictions or less efficient non-convex techniques, neither of which is ideal. (Jain and Kar, 2017) Since many objects in the real world are non-convex, we want to be able to determine how convex an object is, which we can use with the indices of (non)convexity.

In this section, we will compare the convexity of apples and lemons. Although this may appear to be a contrived example, it could have useful applications in AI harvesting technology. Between the two fruits, there are differing areas of concavity and convexity. By assessing the surface of the object, we can determine the type of fruit it is and whether it should be collected at all.

We are not able to work with real lemons or apples (as this is a mathematical paper) but we are able to find representations of these objects in mathematical form. They will serve as a theory that can be applied to the real world.

To start, we define some important terms:

**Definition 8.1.** A set $C$ is convex if given any two points $x$ and $y$ in $C$, the entire line segment that joins $x$ and $y$ is contained within $C$.

Definitions below are sourced from Wolfram MathWorld and reworded to help my own understanding.

**Definition 8.2.** A lemon is a surface of revolution that consists of less than half of a circular arc rotated about a central axis. We will use the upper half of the lemon for this proof. This equation for the upper volume is

$$z \leq \sqrt{R^2 - (\sqrt{x^2 + y^2} + r)^2}$$

where $R > r$.

**Definition 8.3.** An apple is a surface of revolution that consists of a more than half of a circular arc rotated about a central axis. We will use the upper half of the apple for this proof. This equation for this upper volume is

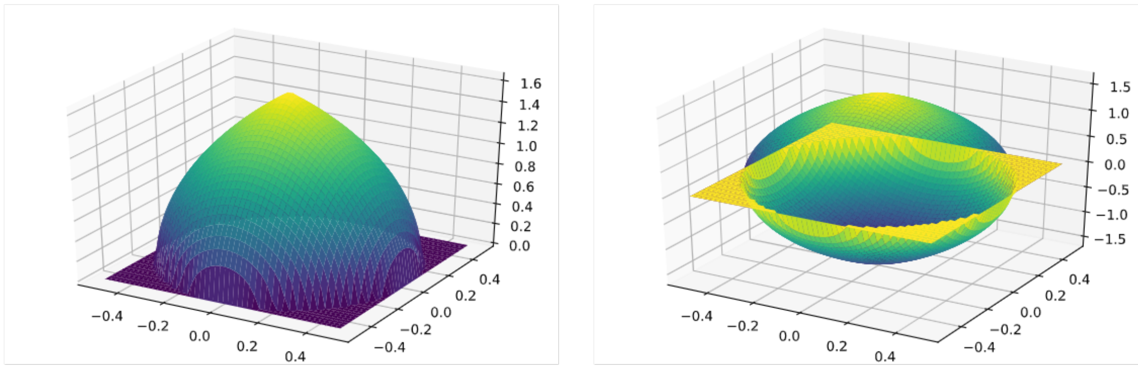
$$z \leq \sqrt{R^2 - (\sqrt{x^2 + y^2} - r)^2}$$

where $R > r$.

The apple and lemon objects appear to be similar from the equation, however, it is the difference between the $+$ and $-$ that causes one fruit to have some convex areas that are not found in the other.

## 8.1 Lemon Convexity

The lemon is a mathematical object defined by a surface of revolution, representing an almost ideal version of the physical fruit. Since the fruit is symmetrical, we will only be working with one half of the lemon and treat the other side as 'stuck in the ground', see Figure 8.1.




(a) Upper Lemon (b) Full Lemon

Figure 8.1: The upper lemon and full lemon.

First we will show that the upper half of the lemon

$$L(0) := \{(x_1, x_2, x_3) \in R^2 : x_3 \leq \sqrt{R^2 - (\sqrt{x_1^2 + x_2^2} + r)^2}\}$$

(centered at $(0, 0, 0)$) is convex. That is, we need to show that if $\mathbf{x}, \mathbf{y} \in L(0)$ then

$$\mathbf{z} := t\mathbf{x} + (1-t)\mathbf{y} \in L(0)$$

for every $t \in [0, 1]$. Explicitly, this means that we need to show that if

$$x_3 \leq \sqrt{R^2 - (\sqrt{x_1^2 + x_2^2} + r)^2}$$

and

$$y_3 \leq \sqrt{R^2 - (\sqrt{y_1^2 + y_2^2} + r)^2},$$

then

$$z_3 \leq \sqrt{R^2 - (\sqrt{z_1^2 + z_2^2} + r)^2},$$

where $z_1 = tx_1 + (1-t)y_1$, $z_2 = tx_2 + (1-t)y_2$ and $z_3 = tx_3 + (1-t)y_3$ are the coordinates of $\mathbf{z} = (z_1, z_2, z_3)$. Since $\mathbf{x}, \mathbf{y} \in L(0)$ and both sides are positive, we have:

$$x_3^2 + (\sqrt{x_1^2 + x_2^2} + r)^2 \leq R^2,$$

$$tx_3^2 + t(\sqrt{x_1^2 + x_2^2} + r)^2 \leq tR^2,$$

$$y_3^2 + (\sqrt{y_1^2 + y_2^2} + r)^2 \leq R^2,$$

$$(1-t)y_3^2 + (1-t)(\sqrt{y_1^2+y_2^2}+r)^2 \leq (1-t)R^2,$$

Adding the equations together

$$tx_3^2 + (1-t)y_3^2 + t(\sqrt{x_1^2+x_2^2}+r)^2 + (1-t)(\sqrt{y_1^2+y_2^2}+r)^2 \leq tR^2 + (1-t)R^2,$$

we use the fact that $f(x) = x^2$ is convex:

$$(tx_3 + (1-t)y_3)^2 + (t\sqrt{x_1^2+x_2^2} + (1-t)\sqrt{y_1^2+y_2^2} + tr + (1-t)r)^2 \leq tR^2 + R^2 - tR^2,$$

and that $f(\mathbf{x}) = \sqrt{x_1^2+x_2^2}$ is convex:

$$(tx_3 + (1-t)y_3)^2 + (\sqrt{(tx_1+(1-t)y_1)^2 + (tx_2+(1-t)y_2)^2} + r)^2 \leq R^2.$$

This is equivalent to

$$(z_3)^2 + (\sqrt{z_1^2+z_2^2}+r)^2 \leq R^2,$$

which is true because the bounds of $\sqrt{z_1^2+z_2^2}+r$ are $[-R+2r, R]$ (from the definition of a lemon). This definitely satisfies the inequality and can be rearranged as follows:

$$z_3 \leq \sqrt{R^2 - (\sqrt{z_1^2+z_2^2}+r)^2}$$

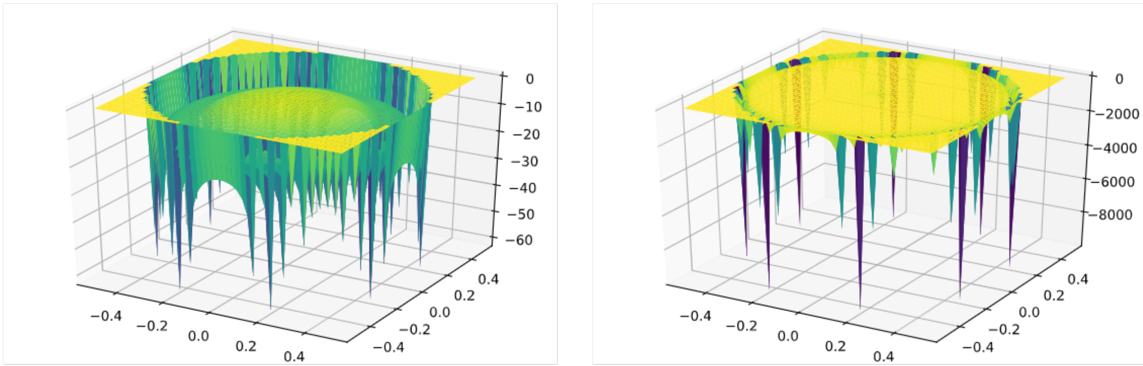
proving that the upper half of the lemon $L(0)$ is convex.

We will apply the indices of convexity and non-convexity to the surface of $L(0)$, or equivalently,

$$x_3 = \sqrt{R^2 - (\sqrt{x_1^2+x_2^2}+r)^2}.$$

This will let us determine the 'level' of convexity at different parts of the surface. The $2 \times 2$ Hessian was computed for the function and its positive and negative eigenvalues were plotted in Figure 8.2. Notice that the plotted values are quite negative, which is a foreshadow to the level of convexity/concavity of the surface. These numbers can be inserted into the formulas for indices of convexity and non-convexity provided by Davydov, Moldavskaya, and Zitikis (2019). In Figure 8.3, the index of lack of convexity is plotted and it can be seen that the entire lemon has an index of 1. This indicates that the surface of the top half of a lemon is non-convex. Similarly, in Figure 8.4, the index of convexity has a value of 0 for the entire lemon.

As can be seen, the surface of the top half of a lemon is quite clearly non-convex, which is somewhat expected with a visual inspection. This defining property of a lemon can be used when harvesting to pick up the fruit. A further inspection of the volume of the fruit can verify that it is in fact a lemon, not something else, like a tennis ball.




(a) Positive (b) Negative

Figure 8.2: Eigenvalues of the $2 \times 2$ Hessian, $\lambda_1$ and $\lambda_2$

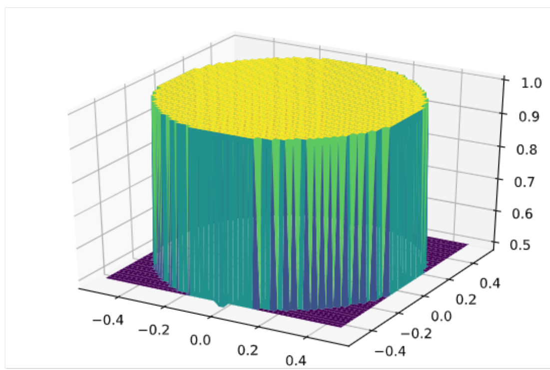



Figure 8.3: The normalized index of lack of convexity $\text{NLOC}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^-(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

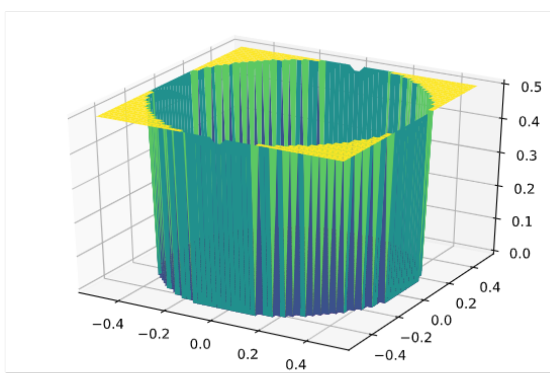
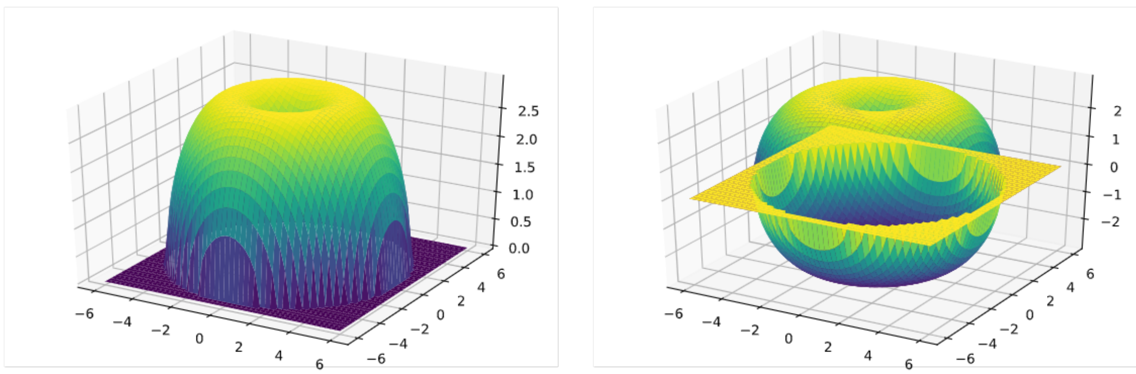



Figure 8.4: The index of convexity $\text{CONV}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^+(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

## 8.2 Apple Non-Convexity

Similar to a lemon, the apple is a mathematical object defined by a surface of revolution. Some say it is an idealized version of the fruit, other say it looks more like a plum. It is also symmetrical and we will only work with the top half of the fruit because when harvesting, the other half will have fallen into the ground. As seen in Figure 8.5, there is a small 'dent' in the center of the fruit which is what separates it from a lemon. This can be used to differentiate the fruits during harvest.




(a) Upper Apple

(b) Full Apple

Figure 8.5: The upper apple and full apple.

We will prove that the upper half of an apple

$$A(0) := \{(x_1, x_2, x_3) \in R^2 : x_3 \leq \sqrt{R^2 - (\sqrt{x_1^2 + x_2^2} - r)^2}\}$$

is not convex in a similar manner to that of the lemon. If the apple were to be convex, we want to find points $\mathbf{x}$ and $\mathbf{y}$ such that there exist a point $t \in [0, 1]$ such that $\mathbf{z} := t\mathbf{x} + (1 - t)\mathbf{y} \in A(0)$. Explicitly, this means that we need to show that if

$$x_3 \leq \sqrt{R^2 - (\sqrt{x_1^2 + x_2^2} - r)^2}$$

and

$$y_3 \leq \sqrt{R^2 - (\sqrt{y_1^2 + y_2^2} - r)^2},$$

then

$$z_3 \leq \sqrt{R^2 - (\sqrt{z_1^2 + z_2^2} - r)^2},$$

where $z_1 = tx_1 + (1-t)y_1$, $z_2 = tx_2 + (1-t)y_2$ and $z_3 = tx_3 + (1-t)y_3$ are the coordinates of $\mathbf{z} = (z_1, z_2, z_3)$. Since $\mathbf{x}, \mathbf{y} \in A(0)$ and both sides are positive, we have:

$$x_3^2 + (\sqrt{x_1^2 + x_2^2} - r)^2 \leq R^2,$$

$$tx_3^2 + t(\sqrt{x_1^2 + x_2^2} - r)^2 \leq tR^2,$$

$$y_3^2 + (\sqrt{y_1^2 + y_2^2} - r)^2 \leq R^2,$$

$$(1-t)y_3^2 + (1-t)(\sqrt{y_1^2 - y_2^2} - r)^2 \leq (1-t)R^2,$$

Adding the equations together

$$tx_3^2 + (1-t)y_3^2 + t(\sqrt{x_1^2 + x_2^2} - r)^2 + (1-t)(\sqrt{y_1^2 + y_2^2} - r)^2 \leq tR^2 + (1-t)R^2,$$

we use the fact that $f(x) = x^2$ is convex:

$$(tx_3 + (1-t)y_3)^2 + (t\sqrt{x_1^2 + x_2^2} + (1-t)\sqrt{y_1^2 + y_2^2} - tr - (1-t)r)^2 \leq tR^2 + R^2 - tR^2,$$

and that $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$ is convex:

$$(tx_3 + (1-t)y_3)^2 + (\sqrt{(tx_1 + (1-t)y_1)^2 + (tx_2 + (1-t)y_2)^2} - r)^2 \leq R^2.$$

This is equivalent to

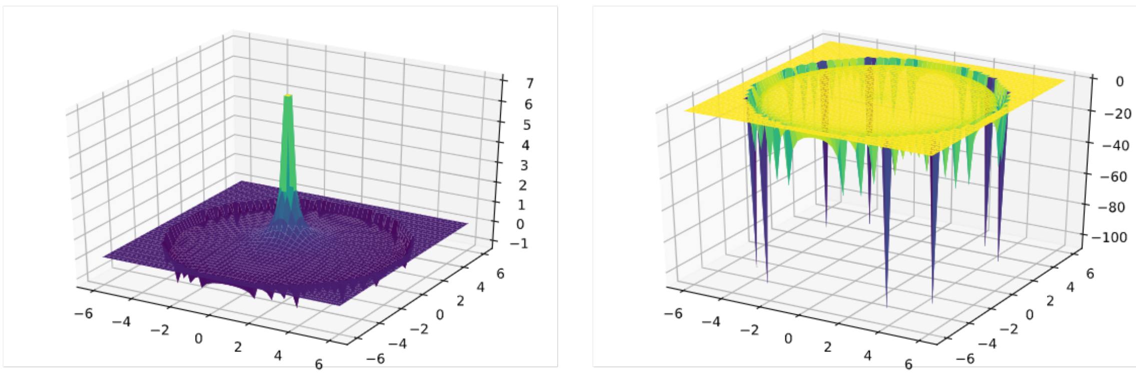
$$(z_3)^2 + (\sqrt{z_1^2 + z_2^2} - r)^2 \leq R^2,$$

which is *not* always true because the bounds of $\sqrt{z_1^2 + z_2^2} + r$ are $[-R - 2r, R]$. The lower bound squared does not satisfy the inequality and proves that the upper half of the apple $A(0)$ is not convex.

We will apply the indices of convexity and non-convexity to the surface of $A(0)$

$$x_3 = \sqrt{R^2 - (\sqrt{x_1^2 + x_2^2} - r)^2}.$$

First, the $2 \times 2$ Hessian was computed and its positive and negative eigenvalues were plotted in Figure 8.6. Note that unlike the lemon, there are some extremely positive values in the center of the apple. But otherwise, it is similar to the lemon and this is enough information to plot the normalized index of lack of convexity (Figure 8.7) and index of convexity (Figure 8.8). The rendered plots are notable in that they have clear dips and spikes, which indicate that the different regions of the graph have different levels of convexity and non-convexity. Earlier, the 'dent' in the apple was mentioned and the properties of that 'dent' are clearly seen here: as a dip in the NLOC and as a spike in CONV. Based on those findings, the 'dent' is an area of convexity that is unlike the other concave areas of the apple. AI harvesting technology can identify these areas and classify this type of object as an 'apple', or 'plum', whichever the operator pleases.




(a) Positive (b) Negative

Figure 8.6: Eigenvalues of the $2 \times 2$ Hessian, $\lambda_1$ and $\lambda_2$

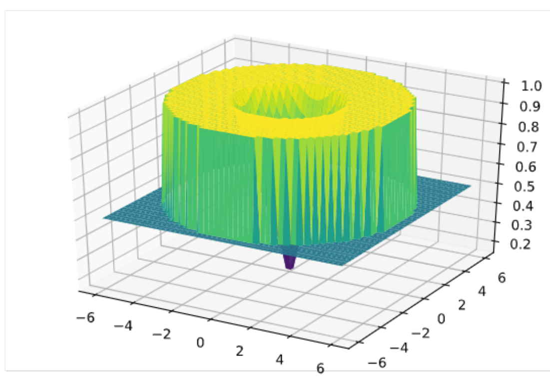



Figure 8.7: The normalized index of lack of convexity $\mathrm{NLOC}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^-(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

Using these properties of the different fruits, it can be made quite clear the distinction between lemons and apples. Indices of convexity/non-convexity show the difference between the lemon with perfect concavity and the apple with a convex dent. This can be applied to future AI harvesting technology as a reliable classifier that does not depend on light or colour. There may even be the possibility of direct

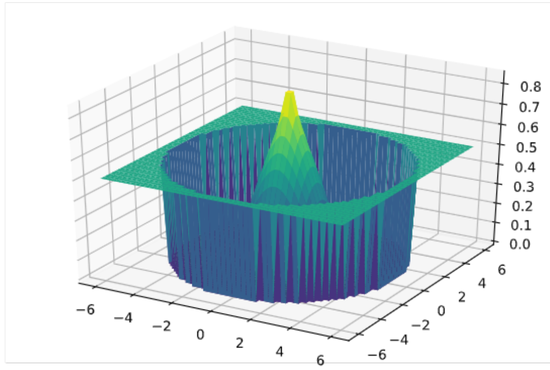



Figure 8.8: The index of convexity $\text{CONV}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^+(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

analysis of the resulting graph through the Hessian and second-derivative which could allow for automatic determination of the maximums and minimums to separate the fruits. However, in the real-world the fruits are not perfect and thus the resulting graphs would not be as flawless either. A tolerance range is necessary to reconcile the differences and provide a best guess as to the type of fruit that is being harvested.

# 9 Discretization

In the real world, there are (obviously) no such perfect objects like the mathematical apple or mathematical lemon. In fact, the challenge to classify these objects is a little bit bigger because computers can only analyze images and cannot observe 3D models. Perhaps in the future this will become possible, but current constraints (technological and budgetary) limit our choice of input.

Thus, it is necessary to perform a discretization of the models that we have worked on. That is, the same computations are performed but this time using numerical rather than symbolic methods.

The discretization was first conducted on the multivariable function $f(x, y) = \sin(xy)$ that was first analyzed. The results are similar to the lemon and apple below, so for brevity, it is not included in the paper.

Below, in Figure 9.1 is the same lemon as earlier but plotted numerically. To avoid errors with 0 and infinity, the edges near $z = 0$ were not plotted.

Using a the numerical values and the lemon function, the Hessian was computed and its eigenvalues plotted in Figure 9.2. The plotted figure is slightly choppier than

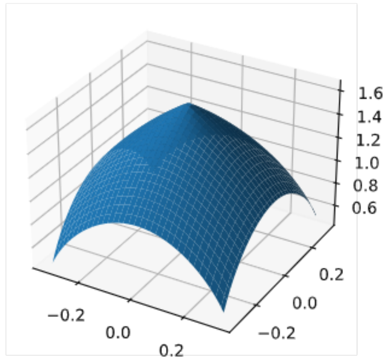
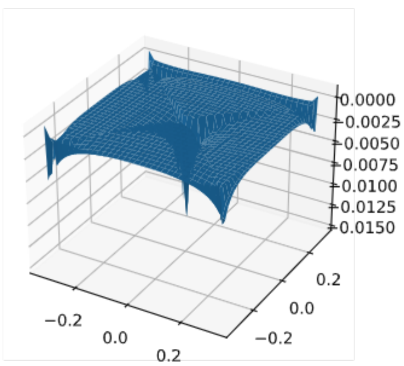
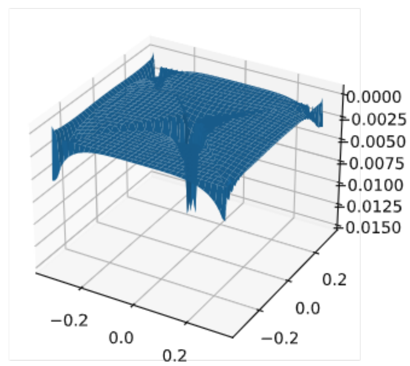



Figure 9.1: The upper lemon, discretized.

the symbolic function.




(a) Positive




(b) Negative

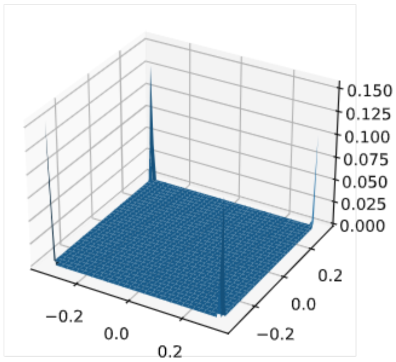
Figure 9.2: Eigenvalues of the $2 \times 2$ Hessian, $\lambda_1$ and $\lambda_2$

Finally, the indices of convexity and non-convexity were computed and plotted once again in Figure 9.3. In this case, there is nothing particularly special about the surface of the lemon as it is non-convex across its entire upper surface.

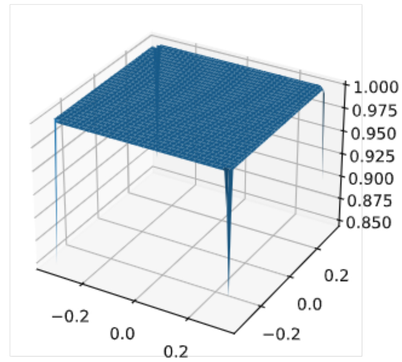
The apple's discretization was performed in a similar manner in Figure 9.4. Edges were not plotted to avoid errors around 0 and infinity.

Using a the numerical values and the apple function, the Hessian was computed and its eigenvalues plotted in Figure 9.5. The plotted figure is slightly choppier than the symbolic function, and the edges are missing due to error avoidance bounds.

Finally, the indices of convexity and non-convexity were computed and plotted once again in Figure 9.6. The results are similar but differences occur in the centre,




(a) $\text{NLOC}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^-(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$




(b) $\text{CONV}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^+(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

Figure 9.3: The normalized index of lack of convexity and the normalized index of convexity for the upper half of a lemon.

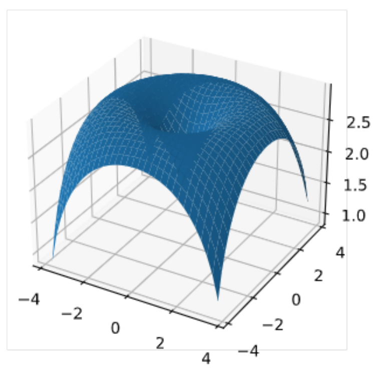
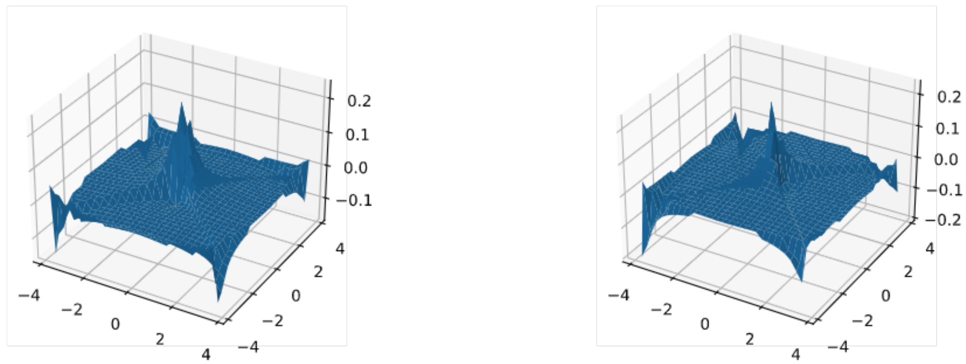



Figure 9.4: The upper apple, discretized.

where the apple has a convex region.

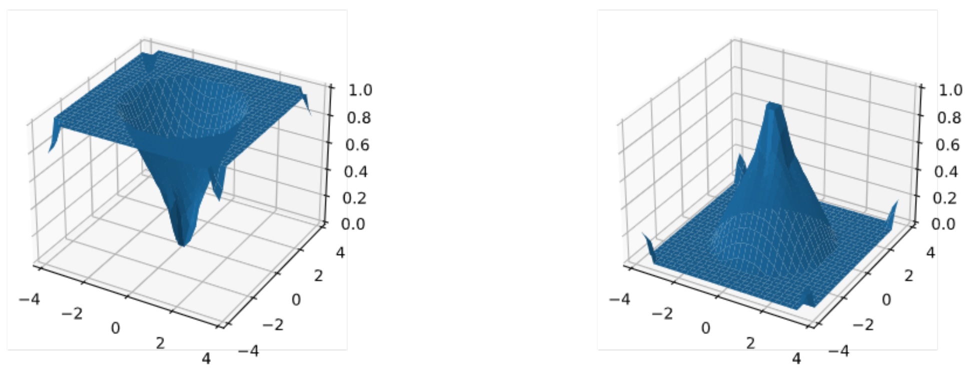
The discretization of both objects is crucially useful because of the way images are stored in pixels. AI will need to analyze the cloud of pixels and map it onto a model, and this discretization allows for easier conversion of pixels to mathematical model. Improvements to the current plots could include a lower spacing to achieve more precise graphs, however, this would greater computational resources. A real-world application would have to take such trade-offs into account as well.




(a) Positive

(b) Negative

Figure 9.5: Eigenvalues of the $2 \times 2$ Hessian, $\lambda_1$ and $\lambda_2$




(a) $\text{NLOC}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^{-}(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

(b) $\text{CONV}(f, \mathbf{x}) = \frac{\sum_{i=1}^{d} \lambda_i^{+}(\mathbf{x})}{\sum_{i=1}^{d} |\lambda_i(\mathbf{x})|}$

Figure 9.6: The normalized index of lack of convexity and the normalized index of convexity for the upper half of a apple.

# 10 Conclusion

In this file, a method of quantifying convexity was developed. Using fundamental concepts from linear algebra, simple multivariate functions were quantified and visualized. Then, the method was extended to look at spherical objects, particularly balls which are 'solid', from which convexity could be determined. These results were then applied to lemons and apples which have broad applications to farming and harvesting techniques. Finally, a discretization of the model was performed which will assist greatly with the implementation on a computer, speeding up computations significantly.

The quantification of convexity is a critical and useful method that has wide-ranging applications. Although only one application has been explored in this working file, uses range from risk to machine learning to other optimization problems. Next steps would be to attach real-world data sources and apply the technique numerically. Some challenges that may need to be overcome include reducing the expensive computations and making it more applicable for complex functions. These changes would make the method available for broader use.

# References


Davydov, Y., Moldavskaya, E., & Zitikis, R. (2019) Searching for and quantifying nonconvexity regions of functions. *Lithuanian Mathematical Journal*, 59, 4, 507–518

Jain, P., & Kar, P. (2017). Non-convex optimization for machine learning. *Foundations and Trends in Machine Learning*, 10, 3-4, 142–336

Kelman, E., & Linker, R. (2014). Vision-based localisation of mature apples in tree images using convexity. *Biosystems Engineering*, 118, 174–185.

Lukacs, E. (1960). *Characteristic Functions.* Hafner Publishing Company, New York.