Gradient Descent for High Dimensional Systems
Lab 1 versus Lab 2

1D Geometry Optimization vs. High Dimensional Optimization

**Methods:**
- Implemented Equations for optimizer

**Applications:**
- Used software to optimize high dimensional systems; perform calculations to predict properties of materials

Today we will fill in the math we left out of how gradient descent will work for high dimensional systems. After lecture, you will apply the content to a worksheet.
1D Potential Energy Surface (PES)

<table>
<thead>
<tr>
<th>Diatomic Molecule (H₂, O₂, etc.)</th>
<th>Two Noble Atoms (Ar₂, etc.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram of 1D PES for Diatomic Molecule" /></td>
<td><img src="image2" alt="Diagram of 1D PES for Two Noble Atoms" /></td>
</tr>
</tbody>
</table>

- Energy required to break a bond
- Optimal Bond length

The majority of chemical systems will need more than one dimension to describe its structure!
2D Potential Energy Surface (PES)

We can view 2D PES by utilizing contour plots.

Contour plots contain contour lines. Contour lines for a function with two variables are lines in which the function has a constant value.
2D Potential Energy Surface (PES)

Another example

You will do a few exercises with contour plots in a worksheet after lecture today.
Potential Energy Surface: a diatomic using *Internal Coordinates*

What is the most probable configuration of the diatomic shown above?

*Geometry Optimization*

*Finding a critical point or local minima on the PES*
Potential Energy Surface: a diatomic using *Internal Coordinates*

What does the Force mean when using Internal Coordinates? 
*If the bond is compressed the force will make the bond length longer and the sign is positive*
Potential Energy Surface: a diatomic using *Internal Coordinates*

What does the Force mean when using Internal Coordinates? 

*If the bond is stretched the force will make the bond length shorter and the sign is negative*
Cartesian Coordinates

Positions of your molecular systems

\[ x_1, y_1, z_1 \quad \text{and} \quad x_2, y_2, z_2 \]

Forces on each atom in your system

\[
\vec{F}_1 = \left\langle F_{x_1}, F_{y_1}, F_{z_1} \right\rangle \\
\vec{F}_2 = \left\langle F_{x_2}, F_{y_2}, F_{z_2} \right\rangle
\]
Potential Energy Surface: a diatomic using *Cartesian Coordinates*

What does the Force mean when using Internal Coordinates?

*If the bond is compressed the force will make the bond length longer*
Potential Energy Surface: a diatomic using *Internal Coordinates*

What does the Force mean when using Internal Coordinates?

*If the bond is stretched the force will make the bond length shorter and the sign is negative*
Force

Given a 1D potential energy surface,

\[ V(r) \]

The force is,

\[ \vec{F}(r) = -\frac{d}{dr} V(r) \]
Force

Given a 2D potential energy surface, \( V(x, y) \)

The force is,

\[
\vec{F}(r) = -\left\langle \frac{\partial}{\partial x} V(x, y), \frac{\partial}{\partial y} V(x, y) \right\rangle
\]

Where \( \frac{\partial}{\partial x} \) is the partial derivative with respect to the variable \( x \)

A partial derivative is the derivative with respect to only one variable when your function is high dimensional.
Force

Given a 2D potential energy surface, $V(x, y)$

The force is, \[ \vec{F}(r) = \left\langle \frac{\partial}{\partial x} V(x, y), \frac{\partial}{\partial y} V(x, y) \right\rangle = -\nabla V(x, y) \]

Where $\frac{\partial}{\partial x}$ is the partial derivative with respect to the variable $x$

A partial derivative is the derivative with respect to only one variable when your function is high dimensional.
Gradient

A gradient is a derivative for a function which has more than one variable. The gradient will be vector with dimensions that are equal to the number of input variables in your function.

Given the function:

\[ V(x, y) \]

the gradient of \( V(x,y) \) is,

\[
\nabla V(x, y) = \left\langle \frac{\partial}{\partial x} V, \frac{\partial}{\partial y} V \right\rangle
\]

Where \( \frac{\partial}{\partial x} \) is the partial derivative with respect to the variable \( x \).
Gradient

A gradient is a derivative for a function which has more than one variable. The gradient will be vector with dimensions that are equal to the number of input variables in your function.

Given the function:

\[ V(x, y, z) \]

the gradient of \( V(x,y) \) is,

\[ \nabla V(x, y, z) = \left\langle \frac{\partial}{\partial x} V, \frac{\partial}{\partial y} V, \frac{\partial}{\partial z} V \right\rangle \]

Where \( \frac{\partial}{\partial x} \) is the partial derivative with respect to the variable \( x \)
Potential Energy Surface:
a diatomic using *Cartesian Coordinates*

\[ V(x_1, y_1, z_1, x_2, y_2, z_2) \]

\[ \vec{F} = - \left( \frac{\partial}{\partial x_1} V, \frac{\partial}{\partial y_1} V, \frac{\partial}{\partial z_1} V, \frac{\partial}{\partial x_2} V, \frac{\partial}{\partial y_2} V, \frac{\partial}{\partial z_2} V \right) \]
Review: Vectors

• Vector: a quantity that contains both a magnitude and a direction.

• Given \( \vec{v} = \langle x_1, x_2, x_3, \ldots, x_N \rangle \)

• The magnitude of the \textit{entire} vector, \( v \), is

\[
\| \vec{v} \| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \ldots + x_N^2}
\]

• The normalized vector, \( \hat{v} \), is

\[
\hat{v} = \frac{\vec{v}}{\| \vec{v} \|}
\]
Normalized vectors

• The normalized vector is

\[ \hat{v} = \frac{\vec{v}}{\|\vec{v}\|} \]

• The normalized vector is also known as the unit vector

• The magnitude of a unit vector is always one but is in the same direction of vector, \( \vec{v} \).

• If you want to change the magnitude of vector \( \vec{v} \), or rescale the magnitude to \( M \) use the following:

\[ \vec{v}_M = M\hat{v} \]
1D Algorithm for Gradient Descent

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.
3. If not, move the atoms such that they go towards a critical points

\[ r_{n+1} = r_n + \alpha F(r_n) \]

4. Repeat.
2D Algorithm for Gradient Descent

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.
3. If not, move the atoms such that they go towards a critical points
   \[
   \langle x, y \rangle_{n+1} = \langle x, y \rangle_n + \alpha \langle F_x, F_y \rangle_n
   \]
4. Repeat.
High Dimensional Algorithm for Gradient Descent

The general procedure for numerical geometry optimization is as follows:

1. Calculate the force on all atoms for some configuration of an atomic system.
2. If the magnitude of the force is less than threshold, you have found a critical point! STOP.
3. If not, move the atoms such that they go towards a critical points

\[
\begin{bmatrix}
  x_1 & y_1 & z_1 \\
  x_2 & y_2 & z_2 \\
  \vdots & \vdots & \vdots \\
  x_A & x_A & x_A \\
\end{bmatrix}_{n+1} = \begin{bmatrix}
  x_1 & y_1 & z_1 \\
  x_2 & y_2 & z_2 \\
  \vdots & \vdots & \vdots \\
  x_A & x_A & x_A \\
\end{bmatrix}_n + \alpha \begin{bmatrix}
  F_{x1} & F_{y1} & F_{z1} \\
  F_{x2} & F_{y2} & F_{z2} \\
  \vdots & \vdots & \vdots \\
  F_{xA} & F_{yA} & F_{zA} \\
\end{bmatrix}_n
\]

Where \( A \) is the number of atoms in the chemical system

4. Repeat.
Potential Energy Surface: a diatomic using *Cartesian Coordinates*

\[ V(x_1, y_1, z_1, x_2, y_2, z_2) \]

\[ \vec{F}(r) = -\langle \frac{\partial}{\partial x_1} V, \frac{\partial}{\partial y_1} V, \frac{\partial}{\partial z_1} V, \frac{\partial}{\partial x_2} V, \frac{\partial}{\partial y_2} V, \frac{\partial}{\partial z_2} V \rangle \]
Vector Review from Tuesday 2/20
Magnitude of Vector

\[ \mathbf{v} = \langle 1, 1 \rangle \]

\[ \| \mathbf{v} \| = \sqrt{1^2 + 1^2} = \sqrt{2} \]

Use Pythagorean theorem to calculate the length of the third side of triangle, or magnitude of vector, \( v \).
Direction of Vector: Normalized vector

\[ \vec{v} = \langle 1, 1 \rangle \]

\[ \| \vec{v} \| = \sqrt{1^2 + 1^2} = \sqrt{2} \]

\[ \hat{\vec{v}} = \frac{\vec{v}}{\| \vec{v} \|} = \frac{1}{\sqrt{2}} \langle 1, 1 \rangle \]

\[ \hat{\vec{v}} = \begin{pmatrix} 1 / \sqrt{2} \\ 1 / \sqrt{2} \end{pmatrix} \]

We now have a vector in the same direction, but with a magnitude of 1!
Direction of Vector: Rescale length of vector

\[ \vec{v} = \langle 1, 1 \rangle \]

\[ \hat{v} = \langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \rangle \]

\[ \vec{v}_M = M \hat{v} \]

\[ \vec{v}_{0.2} = 0.2 \langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \rangle \]

\[ \vec{v}_{0.2} = \left\langle \frac{0.2}{\sqrt{2}}, \frac{0.2}{\sqrt{2}} \right\rangle \]

\[ \vec{v}_{0.2} = \langle 0.141, 0.141 \rangle \]
Vector Problem: How to interpret Force vector

\[ \vec{F} = -\left( \frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial y_1}, \frac{\partial V}{\partial x_2}, \frac{\partial V}{\partial y_2} \right) = \langle 1, 1, -1, -1 \rangle \]

\[ \|\vec{F}\| = \]

\[ \hat{F} = \]

\[ \vec{F}_1 = \langle 1, 1 \rangle \]

\[ \vec{F}_2 = \langle -1, -1 \rangle \]
Vector Problem

You can translate vectors! Next we will translate the $F_1$ vector to act on atom 1!
Vector Problem

Atomic positions

You can translate vectors! Next we will translate the $\vec{F}_1$ and $\vec{F}_2$ vectors to act on atom 1 and 2! Notice that the scales are different on these two examples.