CME 216, ME 343 - Spring 2020

Eric Darve, ICME

Brief introduction to machine learning for engineers

In computational engineering, we are used to solving problems by using physical laws.

For example, we may consider a falling body.

We apply Newton's second law to obtain

$$\vec{F} = m\vec{a}$$

The force $\vec{F}$ may then be set equal to $m\vec{g}$ where $\vec{g}$ is gravity.

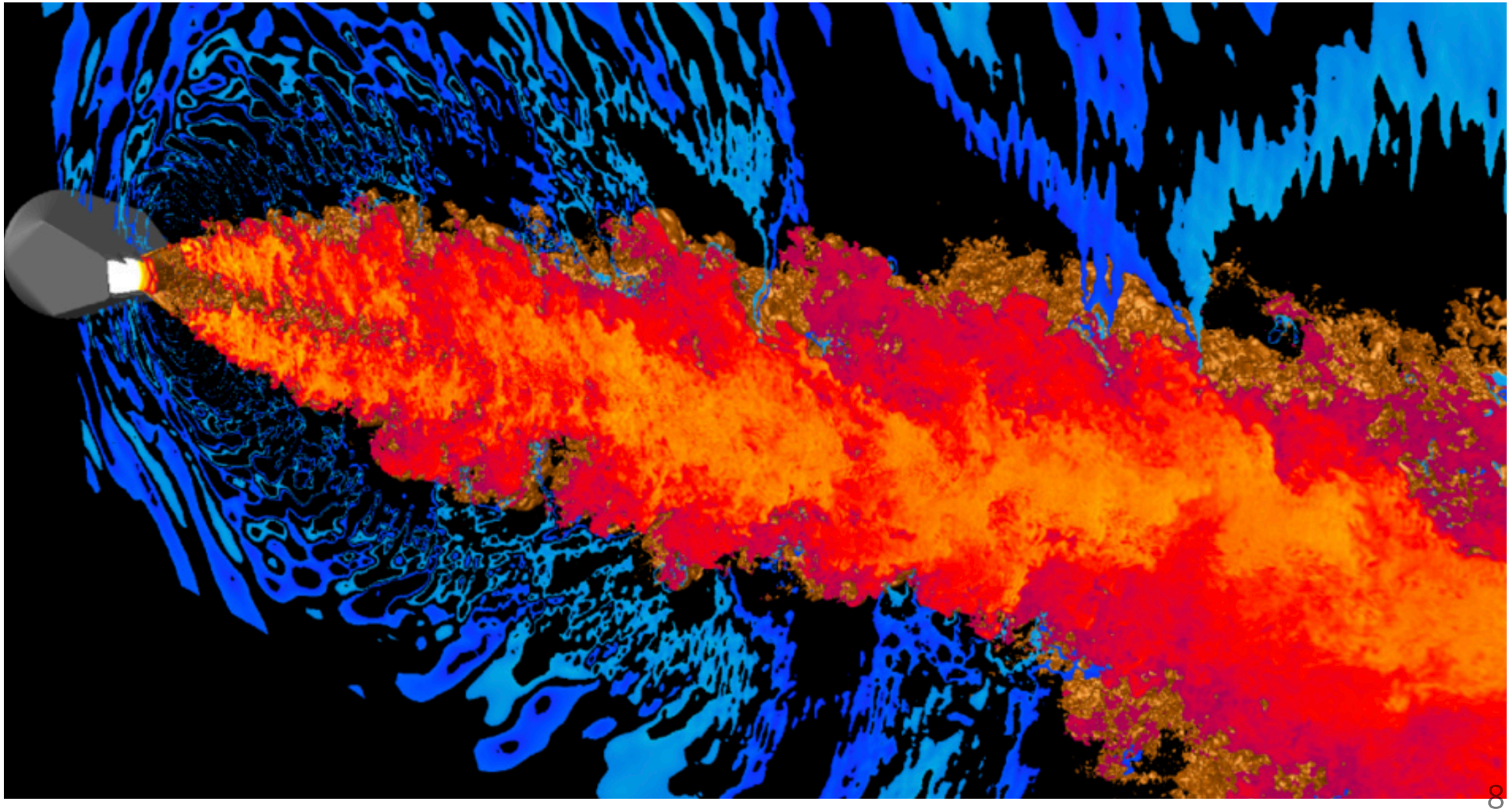
By specifying initial conditions $\vec{r}(0)$ and $\vec{r}'(0)$ and integrating in time, we can find the trajectory $\vec{r}(t)$ for all times t.

Our prediction is based on physical laws and a solution is found after we specify some initial conditions.

There are many other examples such as computing the motion of a spring-mass system based on Hooke's law,

$$F = kx$$

A more complex example is the motion of a fluid flow satisfying the Navier-Stokes equations.



$$\rho \frac{\partial \vec{u}}{\partial t} + \rho \vec{u} \cdot \nabla \vec{u} = -\nabla p + \nabla \cdot \tau + \rho \vec{g}$$

where $\vec{u}$ is the flow velocity, ρ is the density, p the pressure, τ is the deviatoric stress tensor, and $\vec{g}$ represents body forces like gravity.

Machine learning takes a different approach to make predictions.

Generally speaking, it does not make use of physical laws, such as the conservation of momentum or energy (although there are exceptions as we will see), but rather it is based on sampled data, typically called "training data."

A well-known example is the problem of linear regression.

We are given some samples (x_i, y_i), $i = 1, \ldots, m$, and we are interested in building a model of the type $y = w^T x + b$ that relates y to x with a linear relationship.

Typically, the vector w and scalar b do not exist.

That is, there is no w and b such that

$$y_i = w^T x_i + b$$

exactly for all points (x_i, y_i) in our training set.

Instead, for example, we are going to look for w and b such that they minimize the mean square error

$$\sum_{i=1}^{m} |y_i - w^T x_i - b|_2^2$$

More details on linear regression are given in section 5.1.4 of Deep Learning.