Variational Methods

Constrained Maximisation

- Consider a function $f(\boldsymbol{x})$ in three dimensions, and apply a small displacement $\delta\boldsymbol{x} = \delta x + \delta y + \cdots$. **Taylor's Theorem** states that:

  $$f(\boldsymbol{x} + \delta\boldsymbol{x}) = f(\boldsymbol{x}) + \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\partial y + \cdots$$

  So:

  $$\delta \boldsymbol{f} = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\partial y + \cdots = \nabla f \cdot \delta\boldsymbol{x} + \cdots$$

  As the displacement **shrinks to 0**:

  $$\mathrm{d}f = \nabla f \cdot \mathrm{d}\boldsymbol{x}$$

- At an **extremum**, f must be **stationary** – the **first variation** of df must **vanish** for **all directions**. This can only occur if

  $$\boxed{\nabla f = \boldsymbol{0}}$$

- However, if we are trying to maximise **with respect to a constraint** $g(\boldsymbol{x}) = c$, then the first variation in df must also vanish, but **not in all directions**.

  This means that ∇f no longer needs to be 0, but needs to be **perpendicular** to the surface defined by $g(\boldsymbol{x}) = c$.

  However, the normal to the surface $g(\boldsymbol{x}) = c$ is given by ∇g. As such, ∇f needs to be **parallel** to ∇g. In other words, $\nabla f = \lambda \nabla g$. Therefore, we need to solve

  $$\boxed{\begin{gathered}\nabla(f - \lambda g) = \boldsymbol{0} \\ g(\boldsymbol{x}) = c\end{gathered}}$$

  An extension to a higher number of constraints is simple [$\nabla(f - \lambda g - \mu h - \cdots) = \boldsymbol{0}$]

The Euler-Lagrange Equations

- A **functional** is a real-valued mapping whose arguments are functions – ie:

$$F : \text{one or more functions} \to \mathbb{R}$$

- We only consider functionals of the form

$$F[y] = \int_a^b f(x, y, y')\,\mathrm{d}x$$

Where y is a function of x.

- Our task is to find the **form of y which makes stationary our functional** with **fixed values of y** at the **end-points**.

To do this, we consider changing y to some "nearby" function $y(x) + \delta y(x)$ and calculate the corresponding change δF in F.

$$\begin{aligned}
\delta F &= F[y + \delta y] - F[y] \\
&= \int_a^b f(x, y + \delta y, y' + (\delta y)')\,\mathrm{d}x - \int_a^b f(x, y, y')\,\mathrm{d}x \\
&= \int_a^b \overbrace{f(x, y, y') + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}(\delta y)'}^{\text{Taylor expansion, ignoring higher order terms}}\,\mathrm{d}x - \int_a^b f(x, y, y')\,\mathrm{d}x \\
&= \int_a^b \frac{\partial f}{\partial y'}(\delta y)'\,\mathrm{d}x + \int_a^b \frac{\partial f}{\partial y}\delta y\,\mathrm{d}x \\
&= \overbrace{\left[\frac{\partial f}{\partial y'}\delta y\right]_a^b}^{\substack{\text{Must be 0, because} \\ y \text{ is fixed at end-points} \\ \text{and so } \delta y \text{ vanishes at} \\ \text{end-points}}} - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\delta y\,\mathrm{d}x + \int_a^b \frac{\partial f}{\partial y}\delta y\,\mathrm{d}x \\
&= \int_a^b \left\{\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)\right\}\delta y\,\mathrm{d}x
\end{aligned}$$

And since we want $\delta F = 0$ for all possible small variations δy, we must have

$$\boxed{\frac{\partial f}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)}$$

This is the **Euler-Lagrange Equation**.

Note: the *partial* derivatives are **formal** – we evaluate them assuming that y' and y are **unrelated**. However, the complete derivative needs to be done "properly".

- If there are $\boldsymbol{n}$ dependent functions, then the expression above becomes

$$\delta F = \sum_{i=1}^{n} \int_a^b \left\{ \frac{\partial f}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i} \right) \right\} \delta y_i \, \mathrm{d}x$$

Which makes it clear that **every** variable needs to satisfy the **Euler-Lagrange Equation independently**.

- Consider a few simplifying cases:
  - ***f* does not contain *y* explicitely**

    In that case, the Euler-Lagrange equation becomes:

$$\frac{\partial f}{\partial y'} = \text{constant}$$

  - ***f* does not contain *x* explicitly**

    Using the **chain rule** on *f*, we have

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y' + \frac{\partial f}{\partial y'} y''$$

    Then, obtaining $\partial f / \partial y$ from the E-L equation:

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}x} &= \frac{\partial f}{\partial x} + y' \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'} \right) + \frac{\partial f}{\partial y'} y'' \\ &= \frac{\partial f}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x}\left(y' \frac{\partial f}{\partial y'} \right) \end{aligned}$$

    So:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(f - y' \frac{\partial f}{\partial y'} \right) = \frac{\partial f}{\partial x}$$

    If *f* has **no explicit *x* depence**, then

$$f - y' \frac{\partial f}{\partial y'} = \text{constant}$$

    (Or, for n-variables:

$$f - \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'} = \text{constant}\,)$$

Constrained Variation

- Consider a situation in which we want to find the **stationary value** of $F[y]$ **subject to** $G[y] = c$.
- In that case, simply construct the *new* functional

$$F[y] - \lambda G[y]$$

And minimise it, using $G[y] = c$ to **eliminate** $\boldsymbol{\lambda}$.

Physical Examples

- **Fermat's Principle**, in **optics**, states that the path of a **ray of light** will follow a **path** such that the **total optical path length** (*Physical length* × *Refractive index*) is stationary. In other words, **minimising**

$$\int_A^B \mu(\boldsymbol{r})\,\mathrm{d}l$$

- **Hamilton's Principle of Least Action** states that if a **mechanical system** is **uniquely defined** by a **number of coordinates** $\boldsymbol{q_i}$ and **time**, and only experiences forces **derivable from a potential**, then the motion of such a system is such as to make

$$\mathcal{L} = \int_{t_0}^{t_1} L(q_1, \cdots q_n, \dot{q}_1, \cdots \dot{q}_n, t)\,\mathrm{d}t$$

stationary. L is the **Lagrangian** of the system, defined by its **kinetic energy** $\boldsymbol{T}$ and **potential energy** $\boldsymbol{V}$ as:

$$L = T - V$$

Sturm-Liouville Problems – Introduction

- We show that the following three problems are equivalent:
  - Find the **eigenvalues** and **eigenfunctions** that solve the **Sturm-Liouville problem**:

$$-\left[py'\right]' + qy = \lambda \rho y$$

    Between a and b where, in that interval
    - $p(x) \neq 0$
    - $w(x) > 0$

- Find the functions $y(x)$ for which
$$F[y] = \int_a^b [py'^2 + qy^2]\,\mathrm{d}x$$
is **stationary subject to**
$$G[y] = \int_a^b \rho y^2\,\mathrm{d}x = 1$$
- Find the functions $y(x)$ for which
$$\Lambda[y] = \frac{F[y]}{G[y]}$$
is stationary. Λ is the **Rayleigh Quotient**.

Furthermore, if y satisfies the appropriate boundary **constraints** for Sturm-Liouville problems, the value of $F[y]$ and $\Lambda[y]$ are the **eigenvalues** of the problem.

- **EQUIVALENCE OF (1) AND (2)**

  To solve (2), we need to find the **stationary function** of the **functional** $F - \lambda G$. By the E-L equations, this happens where
$$\frac{\mathrm{d}}{\mathrm{d}x}(2py') = 2qy - 2\lambda\rho y$$
$$\boxed{-\frac{\mathrm{d}}{\mathrm{d}x}(py') + qy = \lambda\rho y}$$
  Which is, indeed, the Sturm-Liouville equation. Furthermore, **multiplying by *y*** and **integrating**:
$$\overbrace{\int_a^b -y\frac{\mathrm{d}}{\mathrm{d}x}(py') + qy^2\,\mathrm{d}x = \int_a^b \lambda\rho y^2\,\mathrm{d}x}^{\text{Just the above, multiplied by } y \text{ and integrated}} = \lambda G[y] = \lambda$$
  So:
$$\begin{aligned}\lambda &= -\int_a^b y\frac{\mathrm{d}}{\mathrm{d}x}(py')\,\mathrm{d}x + \int_a^b qy^2\,\mathrm{d}x \\ &= \overbrace{-\left[ypy'\right]_a^b}^{=0\text{ because of BC}} + \int_a^b y'py'\,\mathrm{d}x + \int_a^b qy^2\,\mathrm{d}x \\ &= \int_a^b py'^2 + qy^2\,\mathrm{d}x \\ &= F[y]\end{aligned}$$
  So we do indeed see that the values of $F[y]$ are the eigenvalues.

- **Equivalency of (2) and (3)**

  Consider

$$\begin{aligned}\delta\Lambda &= \frac{F+\delta F}{G+\delta G} - \frac{F}{G} \\ &= \frac{F+\delta F}{G}\left(1+\frac{\delta G}{G}\right)^{-1} - \frac{F}{G} \\ &\approx \frac{F+\delta F}{G}\left(1-\frac{\delta G}{G}\right) - \frac{F}{G} \\ &\approx \frac{\delta F}{G} - \frac{F\delta G}{G^2}\end{aligned}$$

  This means that $\delta\Lambda$ is stationary only if

$$\begin{aligned}\frac{\delta F}{G} &= \frac{F\delta G}{G^2} \\ \delta F - \frac{F\delta G}{G} &= 0 \\ \delta F - \Lambda\delta G &= 0\end{aligned}$$

  The eigenvalue equivalence can be obtained using similar logic.

  Note: in sooth, the first and third methods do not impose $G[y] = 1$, whereas the second method does, but that can easily be fixed by a quick re-scaling of $y \to \alpha y$. This doesn't affect the *linear* equation in (1), nor does it change the *ratio* in (3), so the eigenvalues are still the same.

Sturm-Liouville Problems – Rayleigh-Ritz Method

- Since the eigenvalues λ_i of the Sturm-Liouville equation are the **stationary values** of Λ (*assuming the boundary conditions work*), **any** evaluation of Λ should give a **value** that lies **between** the **highest** and **lowest** eigenvalues of the corresponding Sturm-Liouville equation:

$$\lambda_{\min} \le \Lambda < \lambda_{\max}$$

  One of $\lambda_{\min}$ or $\lambda_{\max}$ will be infinite.

- This allows us to develop a **systematic method** to **estimate** the **lowest eigenvalue** of a Sturm-Liouville Equation:
  - **Re-formulate** the problem as a **variational principle**.
  - Using **whatever clues are available** (eg: symmetry considerations, general Theorems like "the ground state has no nodes", etc…) we make an **educated guess** at the **true eigenfunction with the lowest eigenvalue**.
  - It is **preferable** for the **trial** to have **as many adjustable parameters as possible** (for example, by using **linear combinations** of **trial solutions**).
  - Calculate $\Lambda[y_{\text{trial}}]$, which will **depend** on these **adjustable parameters** – we then calculate the **minimum** of Λ **w.r.t.** these parameters.
  - We can then state that the **lowest eigenvalue** is $\lambda_0 \leq \Lambda_{\min}$.
- The approximation is **good**:
  - If y_{trial} is **close** to the **true** eigenfunction (say **within** $\boldsymbol{O}(\varepsilon)$ of it) then the **calculated** $\Lambda_{\min}$ will be **within** $\boldsymbol{O}(\varepsilon^2)$**.**
  - We can **improve the approximation** by introducing **more adjustable parameters**.
  - If **more** adjustable **parameters fail** to **significantly improve** the **approximation**, we can reasonably be sure that the approximation is **good**.
  - If the **trial solution** contains the **exact** $\boldsymbol{y_0}$ as a **special case**, then $\Lambda_{\min}$ would be **exact**.
- To find **higher eigenvalues**, we simply use **trial solutions** that are **orthogonal** to **all previous trial solutions**.

Sturm-Liouville Problems – Perturbed Operators

- Assume that y_λ is an **eigenfunction** of
$$-(py')' + qy = \lambda \rho y$$
- Now consider a new problem:
$$-(\hat{p}y')' + \hat{q}y = \hat{\lambda}\hat{\rho}y$$
Where
$$\hat{p} = p + \delta p$$
$$\hat{q} = q + \delta q$$
$$\hat{\rho} = \rho + \delta \rho$$
- Now, for this new equation:
$$\begin{aligned}\lambda + \delta\lambda = \hat{\Lambda}(y_\lambda + \delta y) &= \frac{F + \delta F}{G + \delta G} \\ &= (F + \delta F)\frac{1}{G}\left(1 - \frac{\delta G}{G}\right) \\ &= \left(\frac{F}{G} + \frac{\delta F}{G}\right)\left(1 - \frac{\delta G}{G}\right) \\ &= \frac{F}{G} - \frac{F\delta G}{G^2} + \frac{\delta F}{G} \\ &= \lambda - \lambda\frac{\delta G}{G} + \frac{\delta F}{G} \\ &= \lambda + \frac{1}{G}[\delta F - \lambda \delta G]\end{aligned}$$
(To first order).
- However, Λ, and therefore λ is **stationary** with respect to **first-order perturbations** in y_λ. Therefore:
$$\delta\lambda = \frac{\int_a^b (\delta p)y_\lambda'^2 + (\delta q)y_\lambda^2 - \lambda(\delta w)y_\lambda^2 \, \mathrm{d}x}{\int_a^b \rho y_\lambda^2 \, \mathrm{d}x}$$

Practical Tips

- In 3D problems in space, when trying to find maximum values in a certain area and on it, proceed as follows:

  - Solve the problem with the constraint that the points must be **<u>on</u>** the surface of the volume.
  - Solve the problem with no constraint, and just pick up the solutions in the volume.
- To find the **geodesics** on a surface:
  - Find an expression for $\mathrm{d}\boldsymbol{r}$ in an **appropriate coordinate system**.
  - Find an expression for $\int_A^B |\mathrm{d}\boldsymbol{r}|$ in the said coordinate system, **ignoring products of infinitesimal quantities**.
  - Get one of the "d"s out of the square root, to form a **normal integral**.
  - Deduce what A and B are.
- When using **Fermat's Principle**, evaluate $\mathrm{d}l$ as follows:
  - $\mathrm{d}l = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$
  - $\mathrm{d}l = \mathrm{d}x\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}$
  - $\mathrm{d}l = \mathrm{d}x\sqrt{1 + z'^2}$, where $z(x)$ is the path followed by the light.
- In general, it's better **not** to **expand out** expressions in the **functional** – they can be differentiated just fine as-is.
- Notes about **Lagrangians** in **spherical polars**
  - The θ E-L equation will always reveal that

$$r^2\dot{\theta} = \text{constant}$$

    This is the conservation of angular momentum.
- Any variable "**missing**" from the Langrangian implies a **conserved quantity**.