OPTIMISATION

SUNEESH JACOB A.

OPTIMISATION - EXAMPLES

Minimise

$$f(x) = x^3 - 5x^2 + 7x + 1$$

OPTIMISATION - EXAMPLES

Minimise

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

subject to

$$x_1 + 2x_2 \leq 1$$

$$x_1^2 + x_2 \leq 1$$

$$x_1^2 - x_2 \leq 1$$

$$2x_1^2 + x_2 = 1$$

$$0 \leq x_1 \leq 1$$

$$-0.5 \leq x_2 \leq 2$$

WHY OPTIMIZATION?

- Multiple possible solutions
  - Infinitely many
- Satisfaction of constraints

TYPICAL OPTIMISATION PROBLEM

Minimise

$$f(x)$$

subject to

$$g_1(x) \leq 0$$
$$g_2(x) \leq 0$$
$$\vdots$$
$$h_1(x) = 0$$
$$h_2(x) = 0$$
$$\vdots$$

TYPICAL MULTI-OBJECTIVE OPTIMISATION PROBLEM

Minimise

$$f_1(x)$$
$$f_2(x)$$
$$\vdots$$

subject to

$$g_1(x) \leq 0$$
$$g_2(x) \leq 0$$
$$\vdots$$
$$h_1(x) = 0$$
$$h_2(x) = 0$$
$$\vdots$$

METHODS (UNCONSTRAINED)

- Gradient descent method
- L-BFGS method
- Steihaug-Toint method
- Trust-region policy optimisation method
- Nelder-Mead

METHODS (CONSTRAINED)

- Penalty Methods
- Lagrange Method
- Sequential Quadratic Programming
- Augmented Lagrangian

CONDITION FOR OPTIMUM

$$f(x)$$

$$f'(x) = 0$$

$$f''(x) > 0$$

CONDITION FOR OPTIMUM

$$f(x) = x^3 - 5x^2 + 7x + 1$$

$$f'(x) = 0$$

$$f''(x) > 0$$

$$f'(x) = 3x^2 - 10x + 7 = 0 \Rightarrow x = \frac{7}{3}, 1$$

$$f''\left(\frac{7}{3}\right) = 4, \qquad f''(1) = -4$$

CONDITION FOR OPTIMUM

$$f(\boldsymbol{x}) = x^3 - 5x^2 + 7x + 1$$

$$f'(x) = 0$$

$$f''(x) > 0$$

$$f'(x) = 3x^2 - 10x + 7 = 0 \Rightarrow x = \frac{7}{3}, 1$$

$$f''\left(\frac{7}{3}\right) = 4 > 0, \qquad f''(1) = -4 < 0$$

CONDITION FOR OPTIMUM

$$f(x+\delta x) = f(x) + \frac{\delta x}{1!} f'(x) + \frac{\delta x^2}{2!} f''(x) + \cdots$$

$$f(x+\delta x) = f(x) + \frac{\delta x}{1!} f'(x) + \mathcal{O}(\delta x^2)$$

$$f(x+\delta x) - f(x) = \frac{\delta x}{1!} f'(x) + \mathcal{O}(\delta x^2)$$

$$f'(x) = 0$$

CONDITION FOR OPTIMUM

$$f(x+\delta x) = f(x) + \frac{\delta x}{1!} f'(x) + \frac{\delta x^2}{2!} f''(x) + \cdots$$

$$f(x+\delta x) = f(x) + \frac{\delta x}{1!} f'(x) + \frac{\delta x^2}{2!} f''(x) + \mathcal{O}(\delta x^3)$$

$$f(x+\delta x) = f(x) + 0 + \frac{\delta x^2}{2!} f''(x) + \mathcal{O}(\delta x^3)$$

$$f(x+\delta x) - f(x) = \frac{\delta x^2}{2!} f''(x) + \mathcal{O}(\delta x^3)$$

$$f''(x) > 0$$

CONDITION FOR OPTIMUM

Condition for minimum

$$f'(x) = 0$$

$$f''(x) > 0$$

CONDITION FOR OPTIMUM

If $f''(x)$ is 0 then

$$f'''(x) = 0$$

$$f^{iv}(x) > 0$$

CONDITION FOR OPTIMUM

For a multivariate function:

$$f(\boldsymbol{x}+\boldsymbol{\delta x}) = f(\boldsymbol{x}) + \frac{\boldsymbol{\delta x^T}}{1!}\boldsymbol{\nabla_x f} + \frac{1}{2!}\boldsymbol{\delta x^T \nabla_{xx} f \delta x} + \cdots$$

$$\boldsymbol{\nabla_x f = 0}$$

$$\boldsymbol{\nabla_{xx} f} \rightarrow \textbf{positive definite}$$

(All eigenvalues should be greater than zero)

CONDITION FOR OPTIMUM

$$f(\boldsymbol{x}+\boldsymbol{\delta x}) = f(\boldsymbol{x}) + \frac{\boldsymbol{\delta x^T}}{1!}\boldsymbol{\nabla_x f} + \frac{1}{2!}\boldsymbol{\delta x^T \nabla_{xx} f \delta x} + \cdots$$

If $\boldsymbol{x} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}$ then

$$f(\boldsymbol{x}+\boldsymbol{\delta x}) = f(\boldsymbol{x}) + \{\delta x_1 \quad \delta x_2\}\begin{Bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{Bmatrix} + \frac{1}{2}\{\delta x_1 \quad \delta x_2\}\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix}\begin{Bmatrix} \delta x_1 \\ \delta x_2 \end{Bmatrix} + \cdots$$

CONDITION FOR OPTIMUM

$$f(\boldsymbol{x}+\boldsymbol{\delta x}) = f(\boldsymbol{x}) + \{\delta x_1 \quad \delta x_2\}\begin{Bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{Bmatrix} + \frac{1}{2}\{\delta x_1 \quad \delta x_2\}\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix}\begin{Bmatrix} \delta x_1 \\ \delta x_2 \end{Bmatrix} + \cdots$$

$$\begin{Bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{Bmatrix} = 0$$

$$\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix} \rightarrow \text{positive definite}$$

(all eigenvalues should be positive)

LOCAL OPTIMUM VS GLOBAL OPTIMUM

GLOBAL OPTIMUM IS THE MOST OPTIMAL POINT AMONG ALL THE LOCAL OPTIMA

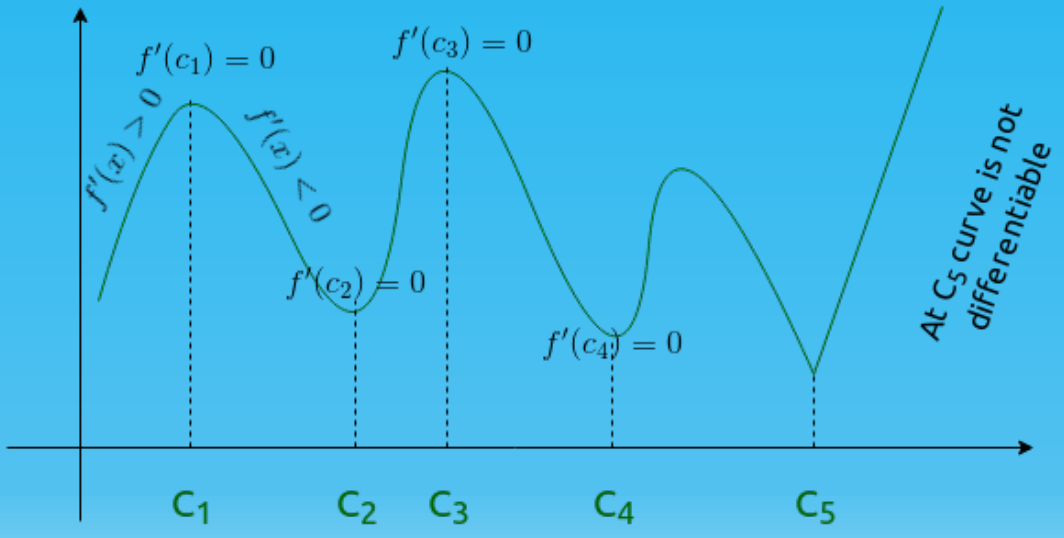

Image taken from: https://media.geeksforgeeks.org/wp-content/uploads/20190520123301/diff1.png

LOCAL OPTIMUM VS GLOBAL OPTIMUM

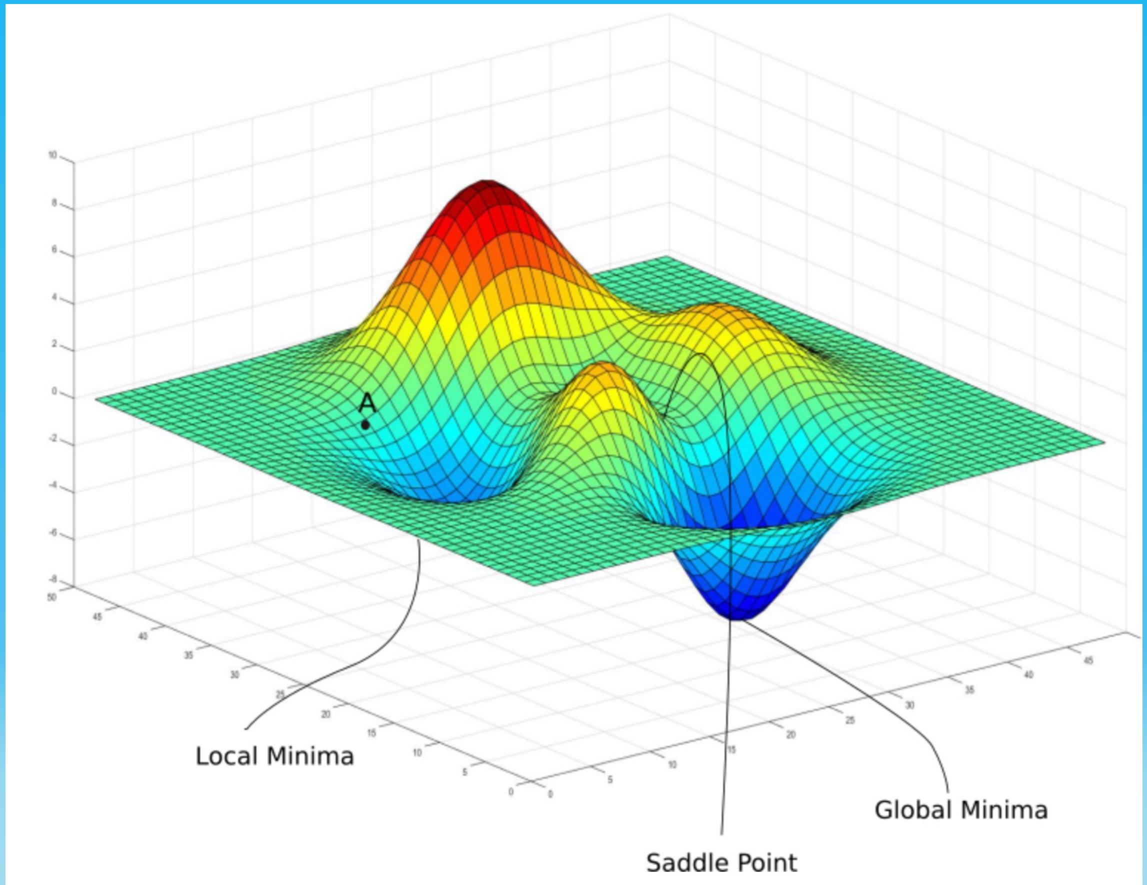

Image taken from: https://wngaw.github.io/images/local_vs_global_minima.png

CONSTRAINED OPTIMISATION

with equality constraints alone

Minimise

$$f(\boldsymbol{x})$$

subject to

$$h_1(\boldsymbol{x}) = 0$$

$$h_2(\boldsymbol{x}) = 0$$

CONSTRAINED OPTIMISATION

with equality constraints alone

Minimise

$$f(x)$$

subject to

$$h_1(x) = 0$$

$$h_2(x) = 0$$

Lagrangian

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \lambda_1 h_1(\boldsymbol{x}) - \lambda_2 h_2(\boldsymbol{x})$$

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x}) = f(\boldsymbol{x}) - H(\boldsymbol{x})$$

where $H(\boldsymbol{x}) = \boldsymbol{\lambda}^T \boldsymbol{h}(\boldsymbol{x})$

$$\boldsymbol{\lambda} = \begin{Bmatrix} \lambda_1 \\ \lambda_2 \end{Bmatrix} \qquad \boldsymbol{h}(\boldsymbol{x}) = \begin{Bmatrix} h_1(\boldsymbol{x}) \\ h_2(\boldsymbol{x}) \end{Bmatrix}$$

CONSTRAINED OPTIMISATION

with equality constraints alone

Lagrangian

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - H(\boldsymbol{x})$$

$$f(\boldsymbol{x} + \boldsymbol{\delta x}) = f(\boldsymbol{x}) + \frac{\boldsymbol{\delta x}^{\boldsymbol{T}}}{1!} \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{f} + \frac{1}{2!} \boldsymbol{\delta x}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{f} \boldsymbol{\delta x} + \cdots$$

$$f(\boldsymbol{x} + \boldsymbol{\delta x}) = f(\boldsymbol{x}) + \frac{\boldsymbol{\delta x}^{\boldsymbol{T}}}{1!} \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{f} + \mathcal{O}(\delta x_i^2)$$

$$f(\boldsymbol{x} + \boldsymbol{\delta x}) = f(\boldsymbol{x}) + \frac{\boldsymbol{\delta x}^{\boldsymbol{T}}}{1!} (\boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{L} + \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{H}) + \mathcal{O}(\delta x_i^2)$$

$$f(\boldsymbol{x} + \boldsymbol{\delta x}) - f(\boldsymbol{x}) = \boldsymbol{\delta x}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{L} + \boldsymbol{\delta x}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{H} + \mathcal{O}(\delta x_i^2)$$

$$\boldsymbol{\nabla}_{\boldsymbol{x}} \boldsymbol{L} = \boldsymbol{0}$$

Also, $\boldsymbol{h}(\boldsymbol{x}) = 0 \Rightarrow \boldsymbol{\nabla}_{\boldsymbol{\lambda}} \boldsymbol{L} = \boldsymbol{0}$

CONSTRAINED OPTIMISATION

with equality constraints alone

Decision variable as a function of parameter t:

$$\boldsymbol{x}_{(t)}$$

Function

$$f(\boldsymbol{x}_{(t)}) = f(\boldsymbol{x}_{(0)}) + \frac{\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2!}\left(\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx} \boldsymbol{f}_{(0)} \dot{\boldsymbol{x}}_{(0)} + \boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)}\right) t^2 + \mathcal{O}(t^3)$$

$$f(\boldsymbol{x}_{(t)}) = f(\boldsymbol{x}_{(0)}) + \frac{\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx} \boldsymbol{f}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3)$$

$$f(\boldsymbol{x}_{(t)}) = f(\boldsymbol{x}_{(0)}) + \frac{\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx} \boldsymbol{f}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\left(\boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^T + \boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T\right) \ddot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3)$$

$$f(\boldsymbol{x}_{(t)}) = f(\boldsymbol{x}_{(0)}) + \frac{\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx} \boldsymbol{f}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3)$$

CONSTRAINED OPTIMISATION

with equality constraints alone

Decision variable as a function of parameter t:

$$\boldsymbol{x}_{(t)}$$

Weighted constraint function

$$H\left(\boldsymbol{x}_{(t)}\right) = H\left(\boldsymbol{x}_{(0)}\right) + \frac{\boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)}t + \frac{1}{2!}\left(\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx}\boldsymbol{H}_{(0)}\dot{\boldsymbol{x}}_{(0)} + \boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)}\right)t^2 + \mathcal{O}(t^3)$$

$$\frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)}t^2 = H\left(\boldsymbol{x}_{(t)}\right) - H\left(\boldsymbol{x}_{(0)}\right) - \frac{\boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)}t - \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \boldsymbol{\nabla}_{xx}\boldsymbol{H}_{(0)}\dot{\boldsymbol{x}}_{(0)}t^2 + \mathcal{O}(t^3)$$

CONSTRAINED OPTIMISATION

with equality constraints alone

Decision variable as a function of parameter t:

$$\boldsymbol{x}_{(t)}$$

Function

$$f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \frac{\nabla_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \nabla_{xx}\boldsymbol{f}_{(0)}\dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\nabla_x \boldsymbol{L}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\nabla_x \boldsymbol{H}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3)$$

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \frac{\nabla_x \boldsymbol{f}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \nabla_{xx}\boldsymbol{f}_{(0)}\dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\nabla_x \boldsymbol{L}_{(0)}^T \ddot{\boldsymbol{x}}_{(0)} t^2 + \left(H\left(\boldsymbol{x}_{(t)}\right) - H\left(\boldsymbol{x}_{(0)}\right) - \frac{\nabla_x \boldsymbol{H}_{(0)}^T}{1!}\dot{\boldsymbol{x}}_{(0)} t - \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^T \nabla_{xx}\boldsymbol{H}_{(0)}\dot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3)\right) + \mathcal{O}(t^3)$$

CONSTRAINED OPTIMISATION

with equality constraints alone

Decision variable as a function of parameter t:

$$\boldsymbol{x}_{(t)}$$

Function

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \frac{\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^{\boldsymbol{T}}}{1!}\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{f}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^{\boldsymbol{T}} \ddot{\boldsymbol{x}}_{(0)} t^2 + \left(H\left(\boldsymbol{x}_{(t)}\right) - H\left(\boldsymbol{x}_{(0)}\right) - \frac{\boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^{\boldsymbol{T}}}{1!}\dot{\boldsymbol{x}}_{(0)} t - \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{H}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \mathcal{O}(t^3) \right) + \mathcal{O}(t^3)$$

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \left(\boldsymbol{\nabla}_x \boldsymbol{f}_{(0)}^{\boldsymbol{T}} - \boldsymbol{\nabla}_x \boldsymbol{H}_{(0)}^{\boldsymbol{T}}\right)\dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}} \left(\boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{f}_{(0)} - \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{H}_{(0)}\right) \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^{\boldsymbol{T}} \ddot{\boldsymbol{x}}_{(0)} t^2 + H\left(\boldsymbol{x}_{(t)}\right) - H\left(\boldsymbol{x}_{(0)}\right) + \mathcal{O}(t^3)$$

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^{\boldsymbol{T}} \dot{\boldsymbol{x}}_{(0)} t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}} \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{L}_{(0)} \dot{\boldsymbol{x}}_{(0)} t^2 + \frac{1}{2}\boldsymbol{\nabla}_x \boldsymbol{L}_{(0)}^{\boldsymbol{T}} \ddot{\boldsymbol{x}}_{(0)} t^2 + H\left(\boldsymbol{x}_{(t)}\right) - H\left(\boldsymbol{x}_{(0)}\right) + \mathcal{O}(t^3)$$

Note: Here, despite $\mathcal{O}(t^3)$ at the end of RHS, the term $H\left(\boldsymbol{x}_{(t)}\right)$ is not limited to second order terms of t but includes all the orders of t !!

CONSTRAINED OPTIMISATION

with equality constraints alone

Let $\boldsymbol{x}_{(t)}$ be always a point in the feasible space. $\therefore\ \boldsymbol{h}\left(\boldsymbol{x}_{(t)}\right) = 0$. This implies that $H\left(\boldsymbol{x}_{(t)}\right)$ is always zero for any t.

$$f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\dot{\boldsymbol{x}}_{(0)}t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}}\boldsymbol{\nabla}_{\boldsymbol{xx}}\boldsymbol{L}_{(0)}\dot{\boldsymbol{x}}_{(0)}t^2 + \frac{1}{2}\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\ddot{\boldsymbol{x}}_{(0)}t^2 + H\left(\boldsymbol{x}_{(0)}\right) - H\left(\boldsymbol{x}_{(t)}\right) + \mathcal{O}(t^3)$$

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\dot{\boldsymbol{x}}_{(0)}t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}}\boldsymbol{\nabla}_{\boldsymbol{xx}}\boldsymbol{L}_{(0)}\dot{\boldsymbol{x}}_{(0)}t^2 + \frac{1}{2}\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\ddot{\boldsymbol{x}}_{(0)}t^2 + \mathcal{O}(t^3)$$

CONSTRAINED OPTIMISATION

with equality constraints alone

$$f\left(\boldsymbol{x}_{(t)}\right) = f\left(\boldsymbol{x}_{(0)}\right) + \boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\dot{\boldsymbol{x}}_{(0)}t + \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}}\boldsymbol{\nabla}_{\boldsymbol{xx}}\boldsymbol{L}_{(0)}\dot{\boldsymbol{x}}_{(0)}t^2 + \frac{1}{2}\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}}\ddot{\boldsymbol{x}}_{(0)}t^2 + \mathcal{O}(t^3)$$

Condition for $\boldsymbol{x}_{(0)}$ to be a local minimum:

We already know that at the local minimum we have $\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L} = 0$, which implies $\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{L}_{(0)}^{\boldsymbol{T}} = 0$.

$$\Rightarrow f\left(\boldsymbol{x}_{(t)}\right) - f\left(\boldsymbol{x}_{(0)}\right) = \frac{1}{2}\dot{\boldsymbol{x}}_{(0)}{}^{\boldsymbol{T}}\boldsymbol{\nabla}_{\boldsymbol{xx}}\boldsymbol{L}_{(0)}\dot{\boldsymbol{x}}_{(0)}t^2 + \mathcal{O}(t^3)$$

$\boldsymbol{\nabla}_{\boldsymbol{xx}}\boldsymbol{L}_{(0)}$ should be positive definite in the feasible space, i.e., in the tangent space, i.e., in the space normal to $\boldsymbol{\nabla}_{\boldsymbol{x}}\boldsymbol{h}$.

CONSTRAINED OPTIMISATION

with equality constraints alone

Minimise

$$f(\boldsymbol{x})$$

subject to

$$h_1(\boldsymbol{x}) = 0$$

$$h_2(\boldsymbol{x}) = 0$$

First order conditions

$$\nabla_{\boldsymbol{x}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = 0$$

$$\nabla_{\boldsymbol{\lambda}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = 0$$

Second order conditions

$$\boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \rightarrow \text{positive definite in tangent plane}$$

$$\boldsymbol{d}^T \boldsymbol{\nabla}_{\boldsymbol{xx}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \boldsymbol{d} > 0 \text{ for all d satisfying } \boldsymbol{d}^T \nabla \boldsymbol{h} = 0$$

CONSTRAINED OPTIMISATION

with equality and inequality constraints

Minimise

$$f(x)$$

subject to

$$h_1(x) = 0$$

$$h_2(x) = 0$$

$$g_1(x) \leq 0$$

$$g_2(x) \leq 0$$

Lagrangian

$$L(x, \lambda) = f(x) + \lambda_1 h_1(x) + \lambda_2 h_2(x) + \mu_1 g_1(x) + \mu_2 g_2(x)$$

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$

CONSTRAINED OPTIMISATION

Minimise

$$f(x)$$

subject to

$$h_1(x) = 0$$

$$h_2(x) = 0$$

$$g_1(x) \leq 0$$

$$g_2(x) \leq 0$$

Rule: Each inequality constraint can be either active or inactive at a given point.

If it is active then it is considered as an equality constraint for that step, and if it is inactive then it is considered as if the constraint does not exist for that step.

Thus, the optimization problem reduces to a problem with either equality constraints alone or an unconstrained problem.

$$\nabla_x L(x) = 0$$

$$\nabla_\lambda L(x) = 0$$

Either $\mu_i = 0$ (inactive) or $g_i = 0$ (active)

$$\nabla_{\mu_i} L(x) = 0 \text{ for all } i \text{ values where } g_i = 0$$

$\mu_i \geq 0$ (signifies the direction of $g_i(x) \leq 0$)

KKT conditions:

$$\nabla_x L(x) = 0$$

$$h_i = 0$$

$$g_i \leq 0$$

$$\mu_i g_i = 0$$

$$\mu_i \geq 0$$

Second-order conditions:

$$\boldsymbol{\nabla_{xx} L(x, \lambda)} \rightarrow \text{positive definite in tangent space with respect to active set of constraints}$$

CONSTRAINED OPTIMISATION

Penalty approach

$$F(x) = f(x) + c \cdot P(x)$$

Minimise

$$F(x)$$

CONSTRAINED OPTIMISATION

Minimise

$$f(x)$$

subject to

$$h_1(x) = 0$$

$$h_2(x) = 0$$

$$P(x) = h_1^2 + h_2^2$$

Minimise

$$F(x) = f(x) + c \cdot P(x)$$

$$F(x) = f + c(h_1^2 + h_2^2)$$

CONSTRAINED OPTIMISATION

Minimise

$$f(x)$$

subject to

$$g_1(x) \leq 0$$

$$g_2(x) \leq 0$$

$$h_1(x) = 0$$

$$P(x) = (\max(0, g_1))^2 + (\max(0, g_2))^2 + h_1^2$$

CONSTRAINED OPTIMISATION

Minimise

$$f(x)$$

subject to

$$g_1(x) \leq 0$$

$$g_2(x) \leq 0$$

$$g_3(x) \leq 5$$

$$g_4(x) \geq 7$$

$$g_5(x) \geq 0$$

$$h_1(x) = 0$$

$$h_2(x) = -8$$

$$P(x) = (\max(0, g_1))^2 + (\max(0, g_2))^2 + (\max(0, g_3 - 5))^2 + (\max(0,7 - g_4))^2 + (\max(0, -g_5))^2 + h_1^2 + (h_2 + 8)^2$$

GRADIENT DESCENT

Minimise

$$f(x)$$

- Start with an initial guess x_0
- Set $x_n = x_0$.
- If $\nabla_x f(x_n) \neq 0$, find the minimum of $f(x_n - \alpha \nabla_x f(x_n))$ and find the corresponding minimiser α. Else terminate with the solution $\mathrm{x}^* = x_n$
- Put $x_n \leftarrow x_n - \alpha \nabla_x f(x_n)$ and repeat the second step until termination.

CONSTRAINED OPTIMISATION

Example

Minimise

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

subject to

$$x_1 + 2x_2 \le 1$$

$$x_1^2 + x_2 \le 1$$

$$x_1^2 - x_2 \le 1$$
$$2x_1^2 + x_2 = 1$$

$$0 \le x_1 \le 1$$

$$-0.5 \le x_2 \le 2$$

Minimise

$$\begin{aligned} &F(x) \\ &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ &+ c\Big((\max(0, x_1 + 2x_2 - 1))^2 + (\max(0, x_1^2 + x_2 - 1))^2 + (\max(0, x_1^2 \\ &- x_2 - 1))^2 + (2x_1^2 + x_2 - 1)^2 + (\max(0, -x_1))^2 \\ &+ (\max(0, x_1 - 1))^2 + (\max(0, -0.5 - x_2))^2 + (\max(0, x_2 - 2))^2\Big) \end{aligned}$$

PYTHON IMPLEMENTATION

```
def F(x,c=0):

    f = 100*(x[1]-x[0]**2)**2+(1-x[0])**2

    g1 = x[0]+2*x[1]-1
    g2 = x[0]**2+x[1]-1
    g3 = x[0]**2-x[1]-1
    h1 = 2*x[0]+x[1]-1
    g4 = -x[0]
    g5 = x[0]-1
    g6 = -0.5-x[1]
    g7 = x[1]-2

    P = max(0,g1)**2+max(0,g2)**2+max(0,g3)**2+max(0,g4)**2+max(0,g5)**2+max(0,g6)**2+max(0,g7)**2+h1**2

    return f+c*P

res = minimize(lambda x:F(x,c=1), [1,1], method='nelder-mead', options={'xatol': 1e-8, 'disp': True})
```

Minimise

$$
\begin{aligned}
&F(x) \\
&= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\
&+ c\Big((\max(0, x_1 + 2x_2 - 1))^2 + (\max(0, x_1^2 + x_2 - 1))^2 + (\max(0, x_1^2 \\
&- x_2 - 1))^2 + (2x_1^2 + x_2 - 1)^2 + (\max(0, -x_1))^2 \\
&+ (\max(0, x_1 - 1))^2 + (\max(0, -0.5 - x_2))^2 + (\max(0, x_2 - 2))^2\Big)
\end{aligned}
$$

DISCRETE VARIABLE OPTIMISATION

$x_1 \rightarrow$ number of chips

The optimized result is $x_1 = 3.219$

Branch and Bound method
Heuristic methods

Example:

Maximise

$$f(x) = 100x_1 + 150x_2$$

subject to

$$2x_1 + x_2 \leq 10$$
$$3x_1 + 6x_2 \leq 40$$
$$x_1, x_2 \geq 0$$

x_1 and x_2 are integers.

GENETIC ALGORITHMS

Minimise

$$f(x) = 3x^2 - 5x + 2$$

Bounds of x

$$[-10,10]$$

$$-10 \rightarrow 0000000000000$$

$$+10 \rightarrow 1111111111111$$

GENETIC ALGORITHMS

- Initial population (Generation 0)
- New generation
  - Fitness function and selection
    - Can be the objective function itself
    - Pairs of individuals
  - Crossover
    - 1001010100101101
    - 0101001010101010
  - Mutation
    - Flipping of a random bit

- 1001010110101010
- 0101001000101101

PYTHON IMPLEMENTATION

```
import numpy as np
from pymoo.algorithms.soo.nonconvex.ga import GA
from pymoo.core.problem import Problem
from pymoo.operators.crossover.sbx import SBX
from pymoo.operators.mutation.pm import PM
from pymoo.operators.repair.rounding import RoundingRepair
from pymoo.operators.sampling.rnd import IntegerRandomSampling
from pymoo.optimize import minimize
```

```
class MyProblem(Problem):

    def __init__(self):
        super().__init__(n_var=2, n_obj=1, n_ieq_constr=4, xl=[-10,-10], xu=[10,10], vtype=int)

    def _evaluate(self, x, out, *args, **kwargs):
        out["F"] = -100*x[:,0]-150*x[:,1]
        out["G"] = [2*x[:,0]+x[:,1]-10, 3*x[:,0]+6*x[:,1]-40,-x[:,0],-x[:,1]]

problem = MyProblem()

method = GA(pop_size=20, sampling=IntegerRandomSampling(), crossover=SBX(prob=1.0, eta=3.0, vtype=float, repair=RoundingRep
mutation=PM(prob=1.0, eta=3.0, vtype=float, repair=RoundingRepair()), eliminate_duplicates=True)

res = minimize(problem, method, termination=('n_gen', 40), seed=1, save_history=True)
```

Queries?

Thank you!