CEE 251L. Uncertainty, Design and Optimization
Department of Civil and Environmental Engineering
Duke University
Homework 3, due: Tuesday, January 27, 2025

This course considers *engineering design* as an application of *engineering analysis* in which the design specification (i.e., the design plan) is quantified by a set of *design variables*

$$\boldsymbol{v} = \left[\; v_1, \quad v_2, \quad v_3, \quad \cdots, \quad v_n \; \right]$$

and through which performance, utility, safety, equity, and sustainability can be quantified and optimized or constrained.

A design that meets performance, utility, safety, equity, and sustainability requirements can be called *functional.* In the approach adopted for this course one of these design requirements is selected to be the primary design objective, and is to be optimized. This is quantified by an *objective function*, which is a function of the design variables.

$$f(v_1, v_2, \cdots, v_n) \qquad \text{or, more compactly,} \qquad f(\boldsymbol{v}) \;,$$

by collecting the n individual design variables into a single vector $\boldsymbol{v}$. A common objective is to simply minimize the total cost of the design, possibly including externalities. In other design problems the objective could be to maximize an aspect of performance.

The functionality of a design depends on more than the primary objective (performance, profits, etc.). Most designs must also meet a number of other criteria (e.g., reliable enough, safe enough, stable enough, strong enough, equitable enough, sustainable enough). These criteria also depend upon the values of the design variables, and are expressed as inequalities. By convention, a set of m inequality constraints can be written as

$$\begin{array}{rcl} g_1(v_1, v_2, v_3, \cdots, v_n) & \leq & 0 \\ g_2(v_1, v_2, v_3, \cdots, v_n) & \leq & 0 \\ & \vdots & \\ g_m(v_1, v_2, v_3, \cdots, v_n) & \leq & 0 \end{array} \qquad \text{or, more compactly,} \qquad \boldsymbol{g}(\boldsymbol{v}) \leq \boldsymbol{0} \;,$$

by collecting the m individual inequalities into a single vector inequality. Design constraints confine the design variables to domains (or sub-spaces) of *admissible* alternatives.

In this framework, the admissible and optimal design $\boldsymbol{v}^*$ minimizes (or maximizes) the primary design objective while satisfying the design constraints:

$$\underset{v_1, v_2, \ldots, v_n}{\text{minimize}} \quad f(\boldsymbol{v}) \qquad \text{such that} \qquad \boldsymbol{g}(\boldsymbol{v}) \leq \boldsymbol{0} \;.$$

In certain rare cases, we can write simple equations for $f(\boldsymbol{v})$ and $\boldsymbol{g}(\boldsymbol{v})$ and use calculus to derive equations for the constrained optimum. In the vast majority of practical problems, however, these equations are much too complicated to be solved with pencil and paper. In such cases computer-aided *analysis* can automate the evaluation of the objective and admissibility of any particular trial design. Further, computer-aided *optimization* allows designers to automatically iterate on candidate designs in order to converge rapidly to an admissible and possibly "optimal" solution. But don't expect too much from optimization. In very challenging design optimization problems it can be computationally impractical to converge to an admissible or feasible design that is perfectly optimal *and* robust to uncertainties.

Sometimes, feasibility suffices; being ok is enough.

1. (5 points) engineering ethics

   Engineering Ethics Question 3

2. (10 points) dimensionality

   Some numerical optimiztation methods are better suited for a particular problem than others. The suitability of the method often depends on the number of design variables (the dimensionality of the problem). Try using `ors`, `nms` and `sqp` to solve the unconstrained optimization problem, $\underset{\boldsymbol{v}}{\text{minimize}} \;\; f(\boldsymbol{v})$

   where

   $$f(\boldsymbol{v}) = \sum_{k=1}^{n} \frac{1}{k} (v_k - k)^2$$

   for $n = 2, 10, 50$ and $-2n \le v_k \le 2n$

   To solve this unconstrained problem within a computational framework for solving constrained problems, simply set `g = np.array([-1])` in order to prevent the constraint from ever binding the solution.

   Discuss your results.

3. (30 points) quadratic programming

   This question involves matrix mathematics. Re-read "A little matrix math" and sections 2 and 4 of "Quadratic Programming and Lagrange Multipliers" for background.

   Consider a quadratic objective function in terms of three variables v_1, v_2, v_3,

   $$f(v_1, v_2, v_3) = 12v_1^2 + 35v_2^2 + 18v_3^2 - 12v_1v_2 + 3v_1v_3 - 18v_2v_3 + 10v_1 - 20v_2 - 26v_3 + 151$$

   (a) Derive three equations for $\partial f/\partial v_1$, $\partial f/\partial v_2$ and $\partial f/\partial v_3$.

   (b) This objective function may be written in a form

   $$f(\boldsymbol{v}) = \frac{1}{2}\boldsymbol{v}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{v} + \boldsymbol{c}^{\mathsf{T}}\boldsymbol{v} + d \tag{1}$$

   where $\boldsymbol{H}$ is a 3×3 Hessian matrix, $\boldsymbol{c}$ is a 3×1 vector and d is a constant. The Hessian matrix represents the curvature of the objective function.

   $$H_{ij} = \frac{\partial^2 f}{\partial v_i \partial v_j} = \frac{\partial}{\partial v_i}\frac{\partial f}{\partial v_j}$$

   Find the numerical values for the *Hessian* matrix $\boldsymbol{H}$ and the vector $\boldsymbol{c}$ corresponding to the objective function $f(\boldsymbol{v})$, above.

   (c) Is $\boldsymbol{H}$ always symmetric ($H_{ij} = H_{ji}$), ($\boldsymbol{H} = \boldsymbol{H}^{\mathsf{T}}$)?

   (d) Show that in this example (as in all quadratic problems) the vector

   $$\begin{bmatrix} \partial f/\partial v_1 \\ \partial f/\partial v_2 \\ \partial f/\partial v_3 \end{bmatrix}$$

   is the same as $\boldsymbol{H}\boldsymbol{v} + \boldsymbol{c}$.

(e) Compute the values of v_1^*, v_2^*, v_3^* that minimize $f([v_1, v_2, v_3])$. This is an unconstrained minimization. What is the value of the objective function $f(\boldsymbol{v}^*)$? You may use `Python` to solve this linear system of three equations and three unknowns.

(f) Suppose that in addition to minimizing f, the parameter values must also satisfy two inequality constraints:

$$g_1(v_1, v_2, v_3) = -6v_1 + 4v_2 - 2v_3 + 8 \le 0$$

and

$$g_2(v_1, v_2, v_3) = 7v_1 - 5v_2 + 3v_3 - 9 \le 0$$

Find the numerical values of the matrix $\boldsymbol{A}$ (2×3) and the vector $\boldsymbol{b}$ (2×1) representing these two constraints as $\boldsymbol{g}(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v} - \boldsymbol{b} \le \boldsymbol{0}_{2\times 1}$.
Show that $A_{ij} = \partial g_i / \partial v_j$.

(g) Using a column-vector of Lagrange multipliers, $\boldsymbol{\lambda}$ and the augmented objective function and assuming both constraints are binding,

$$f_{\mathsf{A}}(\boldsymbol{v}, \boldsymbol{\lambda}) = f(\boldsymbol{v}) + \boldsymbol{\lambda}^{\mathsf{T}} \boldsymbol{g}(\boldsymbol{v}) = \frac{1}{2} \boldsymbol{v}^{\mathsf{T}} \boldsymbol{H} \boldsymbol{v} + \boldsymbol{c}^{\mathsf{T}} \boldsymbol{v} + d + \boldsymbol{\lambda}^{\mathsf{T}} (\boldsymbol{A}\boldsymbol{v} - \boldsymbol{b}) \qquad (2)$$

show that the two necessary conditions for optimality

$$\left. \frac{\partial f_{\mathsf{A}}}{\partial \boldsymbol{v}} \right|^{\mathsf{T}}_{\substack{\boldsymbol{v} = \boldsymbol{v}^* \\ \boldsymbol{\lambda} = \boldsymbol{\lambda}^*}} = \boldsymbol{0}_{3\times 1} \qquad \text{and} \qquad \left. \frac{\partial f_{\mathsf{A}}}{\partial \boldsymbol{\lambda}} \right|^{\mathsf{T}}_{\substack{\boldsymbol{v} = \boldsymbol{v}^* \\ \boldsymbol{\lambda} = \boldsymbol{\lambda}^*}} = \boldsymbol{0}_{2\times 1}$$

may be expressed in the following KKT matrix equation,

$$\begin{bmatrix} \boldsymbol{H} & \boldsymbol{A}^{\mathsf{T}} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^* \\ \boldsymbol{\lambda}^* \end{bmatrix} = \begin{bmatrix} -\mathbf{c} \\ \boldsymbol{b} \end{bmatrix} .$$

Note that since f_{A} is a scalar

$$f_{\mathsf{A}} = f_{\mathsf{A}}^{\mathsf{T}} = \frac{1}{2} \boldsymbol{v}^{\mathsf{T}} \boldsymbol{H} \boldsymbol{v} + \boldsymbol{v}^{\mathsf{T}} \boldsymbol{c} + d + (\boldsymbol{v}^{\mathsf{T}} \boldsymbol{A}^{\mathsf{T}} - \boldsymbol{b}^{\mathsf{T}}) \boldsymbol{\lambda}$$

(h) Using the numerical values for $\boldsymbol{H}$, $\boldsymbol{c}$, $\boldsymbol{A}$ and $\boldsymbol{b}$, compute values for the optimal parameter vector $\boldsymbol{v}^*$ and the Lagrange multipliers $\boldsymbol{\lambda}^*$. You may use `Python` to solve this linear system of five equations and five unknowns. Confirm that $\boldsymbol{A}\boldsymbol{v}^* - \boldsymbol{b} = \boldsymbol{0}$. Are both Lagrange multipliers positive? What is the value of the objective function $f(\boldsymbol{v}^*_{(\mathrm{h})})$ for which $\boldsymbol{A}\boldsymbol{v}^* - \boldsymbol{b} = \boldsymbol{0}$?

(i) If a constraint $\boldsymbol{g}(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v} - \boldsymbol{b}$ is increased from a value $\boldsymbol{g}$ to a value $\boldsymbol{g} + \delta\boldsymbol{g} = \boldsymbol{A}\boldsymbol{v} - \boldsymbol{b} + \delta\boldsymbol{g}$, this is equivalent to writing $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{b} - \delta\boldsymbol{g}$. So, decreasing $\boldsymbol{b}$ by $\delta\boldsymbol{g}$ increases $\boldsymbol{g}$ by $\delta\boldsymbol{g}$.

Repeat part (h) using $\boldsymbol{b} - [\Delta g_1, 0]^{\mathsf{T}}$ in the KKT system. Use $\Delta g_1 = 0.1$. This increases g_1 (and decreases b_1) by Δg_1. What is the new value of the objective function, $f(\boldsymbol{v}^*_{(\mathrm{i})})$? Compare the difference between $f(\boldsymbol{v}^*_{(\mathrm{h})})$ and $f(\boldsymbol{v}^*_{(\mathrm{i})})$ to the average of the λ_1^* values by confirming that

$$\frac{\Delta f}{\Delta g_1} = \frac{f^*_{(\mathrm{i})} - f^*_{(\mathrm{h})}}{\Delta g_1} \approx \frac{\lambda^*_{1(\mathrm{h})} + \lambda^*_{1(\mathrm{i})}}{2}$$

(j) Repeat part (h) using $\boldsymbol{b} - [0, \Delta g_2]^\mathsf{T}$ in the KKT system. Use $\Delta g_2 = 0.1$. This increases g_2 (and decreases b_2) by Δg_2. What is the new value of the objective function, $f(\boldsymbol{v}^*_{(\mathrm{j})})$? Compare the difference between $f(\boldsymbol{v}^*_{(\mathrm{h})})$ and $f(\boldsymbol{v}^*_{(\mathrm{j})})$ to the average of the λ^*_2 valaues by confirming that

$$\frac{\Delta f}{\Delta g_2} = \frac{f^*_{(\mathrm{j})} - f^*_{(\mathrm{h})}}{\Delta g_2} \approx \frac{\lambda^*_{2(\mathrm{h})} + \lambda^*_{2(\mathrm{j})}}{2}$$

(k) If f has units of € and g_i is a constraint related to the mass of the object (with units of kg), what must be the units of λ_i?

(l) Are you getting convinced that the values of the Lagrange multipliers indicate the sensitivity of the objective function to changes in the constraints?

4. (20 points) fit a constrained function to data

Many constrained optimization problems can be solved in one step by using the KKT equations instead of using numerical optimization. Constrained curve fitting is one such example.

(a) Re-read "Constrained Linear Least Squares"

(b) Download `constrained_least_squares.py` and confirm that it runs without error.

(c) Revise `constrained_least_squares.py` in order to fit the equation

$$\hat{y}(t; \boldsymbol{c}) = c_0 + c_1 t^2 + c_2 \sin(2\pi t) + c_3 \cos(\pi t) + c_4 \exp(-(t^2))$$

to a set of simulated noisy data over the domain $0 \le t \le 10$ such that the initial value $y(t = 0)$, the final value $y(t = 10)$ and the rate of change of the initial value $y'(t = 0)$ are known with certainty, and are therefore constraints to this problem. $\hat{y}(0; \boldsymbol{c}) = 9$, $\hat{y}'(0; \boldsymbol{c}) = 56$, and $\hat{y}(10; \boldsymbol{c}) = 35$.
To generate the measured data, use "true" coefficients
($c_0 = 1$, $c_1 = 0.5$, $c_2 = 4$, $c_3 = 5$, $c_4 = -15$) .

To solve this problem, you need only edit the lines ending with . . . `# *`

(d) Include the numerical results and a plot of figure 1 generated by the program

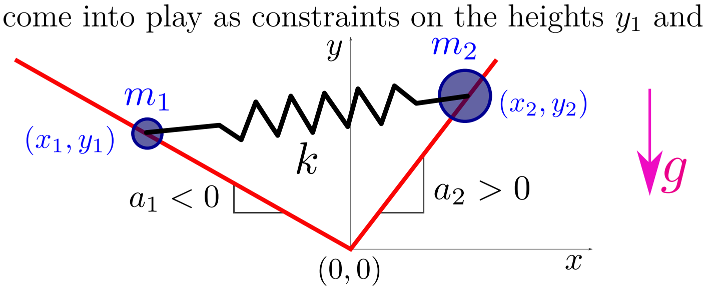
(e) Run `constrained_least_squares.py` a few times.

i. Which coefficient for the unconstrained fit seems most variable?
How variable is it?

ii. Which coefficient for the constrained fit seems most variable?
How variable is it?

5. (30 points) minimize the total potential energy (Π) subject to constraints

Two discs of mass m_1 and m_2 can roll along ramps with slopes $a_1 < 0$ and $a_2 > 0$ and are connected by a spring with spring constant k. The position coordinates of the disk centers are (x_1, y_1) and (x_2, y_2). (For the moment, disregard the existence of the ramps. The ramps will come into play as constraints on the heights y_1 and y_2 of the two disks.)



The total potential energy of the system is the elastic potential energy in the spring $\frac{1}{2}kd^2$ plus the gravitational potential energy of the disks $m_1 g y_1 + m_2 g y_2$. Assuming the unstretched length of the spring is zero, the stretch of the spring is $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

(a) Derive an expression for the total potential energy Π of the system, which is the sum of the elastic potential energy and the gravitational potential energy.

(b) The equilibrium configuration of this system minimizes the total potential energy Defining the coordinate vector $\boldsymbol{q}$ of this system as

$$\boldsymbol{q} = \left[\begin{array}{cccc} x_1, & y_1, & x_2, & y_2 \end{array} \right]^{\mathsf{T}}$$

express

$$\left. \frac{\partial \Pi}{\partial \boldsymbol{q}} \right|_{\boldsymbol{q}=\boldsymbol{q}^*}^{\mathsf{T}} = \boldsymbol{0}_{4\times1} \qquad \text{as} \qquad \boldsymbol{H}\boldsymbol{q}^* + \boldsymbol{c} = \boldsymbol{0}_{4\times1}$$

What is the matrix $\boldsymbol{H}$ in terms of k and what is the vector $\boldsymbol{c}$ in terms of m_1, m_2 and g? What is the rank of $\boldsymbol{H}$? So, is $\boldsymbol{q}^*$ unique? Why does this make sense, physically?

(c) Now consider the constraints $y_1 \geq a_1 x_1$ and $y_2 \geq a_2 x_2$. Write these two constraint equations as $\boldsymbol{A}\boldsymbol{q} - \boldsymbol{b} \leq \boldsymbol{0}_{2\times1}$. What is $\boldsymbol{A}$ in terms of a_1 and a_2? What is $\boldsymbol{b}$?

Adjoining the total potential energy Π with the constraints multiplied by the Lagrange multipliers,

$$\Pi_{\mathsf{A}} = \frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\boldsymbol{H}\boldsymbol{q} + \boldsymbol{\lambda}^{\mathsf{T}}(\boldsymbol{A}\boldsymbol{q} - \boldsymbol{b})$$

This should look very similar to problems 2 and 3.

Without computing any other derivitives write the KKT equations that maximizes Π_{A} in terms of $\boldsymbol{\lambda}$ and minimizes Π_{A} in terms of $\boldsymbol{q}$.

$$\left. \frac{\partial \Pi_{\mathsf{A}}}{\partial \boldsymbol{q}} \right|_{\substack{\boldsymbol{q}=\boldsymbol{q}^* \\ \boldsymbol{\lambda}=\boldsymbol{\lambda}^*}}^{\mathsf{T}} = \boldsymbol{0}_{4\times1} \qquad \text{and} \qquad \left. \frac{\partial \Pi_{\mathsf{A}}}{\partial \boldsymbol{\lambda}} \right|_{\substack{\boldsymbol{q}=\boldsymbol{q}^* \\ \boldsymbol{\lambda}=\boldsymbol{\lambda}^*}}^{\mathsf{T}} = \boldsymbol{0}_{2\times1}$$

(d) Write the KKT matrix equation for this problem in terms of $k, a_1, a_2, m_1, m_2, g, \boldsymbol{q}^*$ and $\boldsymbol{\lambda}^*$. What is the rank of the KKT matrix? Feeling better?

(e) Using numerical values of $m_1 = 3$ kg, $m_2 = 2$ kg, $a_1 = -3$ m/m, $a_2 = 4$ m/m, $k = 20$ N/m and $g = 9.81$ m/s^2, solve the system of six equations and six unknowns for (optimal) equilibrium values of x_1^*, y_1^*, x_2^*, y_2^*, λ_1^* and λ_2^*.

(f) What is the force in the spring?

(g) Does the solution satisfy the constraints $y_1 \geq a_1 x_1$, $y_2 \geq a_2 x_2$? Are both Lagrange multipliers positive? Are both constraints binding? Are you surprised by the values of $\boldsymbol{q}$? Now that you see the numerical values, can you explain why they actually make sense? This is an example of, "Be careful what you ask for, it may not be what you want."

(h) The constraint forces are $-\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\lambda}$. Do these forces act perpendicularly to the ramps?

This method finds the static equilibrium forces *and* the coordinate displacements of elastic systems that are statically determinate or statically indeterminate without the need to derive equilibrium equations (like $\sum F_x = 0$ or $\sum F_y = 0$).
... *cool?* ... you're welcome.