Due date of the homework: **October 7th, midnight**. Submit one PDF together with the source and code files. Show the mathematical work, report solver status (if infeasible), and check every numerical solution against the original constraints. You may use CVX in MATLAB or CVXPY in Python. Up to you.

---

1. Consider the economic dispatch problem

$$\begin{array}{ll} \text{minimize} & 0.025P_{G_1}^2 + 20P_{G_1} + 0.05P_{G_2}^2 + 25P_{G_2} + \pi \\ \text{subject to} & P_{G_1} + P_{G_2} = P_L, \qquad P_{G_1}, P_{G_2} \geq 0. \end{array}$$

   (a) Solve the problem analytically and with CVX/CVXPY when $P_L = 250$ MW.

   (b) Plot both optimal generator outputs for $P_L \in \{150, 200, 250, 300, 350, 400\}$ MW. Use labeled axes, units, markers, a legend, and a grid.

   (c) Explain how changing either quadratic coefficient changes the sharing of load.

   (d) Add the capacity limit $P_{G_1} \leq 160$ MW and resolve the problem at $P_L = 250$ MW.

   (e) Explain, without algebra, what the constant π changes and what it does not change.

2. Prove that $\mathcal{S} = \{x \in \mathbb{R}^n : a^\mathsf{T} x \leq 0\}$ is a convex cone. State carefully which signs the scalar multipliers must have.

3. Prove that the slab $\mathcal{S} = \{x \in \mathbb{R}^n : \alpha \leq a^\mathsf{T} x \leq \beta\}$ is convex, assuming $\alpha \leq \beta$.

4. Show that $\Omega = \{y \in \mathbb{R}^2 : \|y\|_2^2 \leq 4\}$ is convex. Give a norm-based proof.

5. Let

$$A = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix}.$$

   Describe $A \succeq 0$ using scalar polynomial inequalities in $x_1, \ldots, x_6$. Include every principal minor.

6. Find every $\beta \in \mathbb{R}$ for which

$$f(x) = 2x_1^2 + x_2^2 + 3x_3^2 - 2x_1x_3 + 2\beta x_1x_2 + 2x_2x_3$$

   is convex on $\mathbb{R}^3$. Justify necessity and sufficiency.

   Hint: write $f = x^\mathsf{T} Q x$ with $Q = Q^\mathsf{T}$ and test all principal minors of Q.

7. Classify each function as convex, concave, both, or neither on the stated domain:

   (a) $f(x) = \log x$ on $\mathbb{R}_{++}$;

   (b) $g(x_1, x_2) = x_1^2/x_2$ on $\mathbb{R} \times \mathbb{R}_{++}$.

   Use derivatives for the first function. For the second, use its Hessian or recognize it as the perspective of a quadratic.

8. Consider

$$\mathcal{F} = \{x \in \mathbb{R}^2 : x_1 + 2x_2 \geq 2,\ 3x_1 + x_2 \geq 3,\ x_1, x_2 \geq 0\}.$$

   Draw a clear two-dimensional plot of $\mathcal{F}$, including its boundary lines and vertices. For each objective below, give the optimal set and value and identify any unbounded case. Verify the results in CVX/CVXPY.

$$\text{(a) } x_1 + x_2, \quad \text{(b) } -x_1 - x_2, \quad \text{(c) } x_2, \quad \text{(d) } \max\{x_1, x_2\}, \quad \text{(e) } 4x_1^2 + x_2^2.$$

9. Let

$$P = \begin{bmatrix} 6 & 1 & 0 \\ 1 & 5 & 1 \\ 0 & 1 & 4 \end{bmatrix}, \qquad q = \begin{bmatrix} 7.75 \\ -1.25 \\ -7.25 \end{bmatrix}, \qquad r = 2.$$

For $\min\{\frac{1}{2}x^{\mathsf{T}}Px + q^{\mathsf{T}}x + r : -1 \le x_i \le 1\}$:

(a) prove that $x^* = (-1, \frac{1}{4}, 1)^{\mathsf{T}}$ is optimal using the first-order variational inequality;

(b) give a KKT certificate and prove uniqueness;

(c) verify the answer numerically.

Hint: evaluate $Px^* + q$, then use the signs of $x_1 - x_1^*$ and $x_3 - x_3^*$ over the box.

10. Use a complete KKT analysis that we learned in class to find every candidate minimum of

$$\begin{array}{ll} \text{minimize}_{x_1,x_2} & x_1^2 + 2x_2^2 \\ \text{subject to} & x_1 + x_2 \ge 3.5, \\ & x_2 - \frac{1}{2}x_1^2 \ge 1. \end{array}$$

State which candidate is the optimizer and whether it is local, global, and unique.

Hint: write both inequalities as $g_i(x) \le 0$ before forming the Lagrangian, and examine all four active-set cases.

11. Solve

$$\begin{array}{ll} \text{minimize}_{x_1,x_2} & 2x_1 + x_2^2 \\ \text{subject to} & x_1 - x_2 = 4, \\ & 4x_1^2 + x_2^2 \le 25 \end{array}$$

using KKT conditions. Check both intersections of the line and ellipse and explain why only one has a valid inequality multiplier.

12. Newton's method computes the direction $d^{(k)}$ from

$$\nabla^2 f(x^{(k)})d^{(k)} = -\nabla f(x^{(k)}), \qquad x^{(k+1)} = x^{(k)} + \alpha d^{(k)}.$$

The choice $\alpha = 1$ is called a full Newton step; it means that the entire Newton direction is taken. It is not a time step. Apply this method to

$$f(x_1, x_2) = x_1^2 + 1.5x_2^2 - 2x_1 + 3x_2 + 5, \qquad x^{(0)} = (-1, 2)^{\mathsf{T}}.$$

(a) Derive the gradient, Hessian, and Newton direction. Use $\alpha = 1$ and list the iterates until convergence.

(b) Repeat for eight iterations using $\alpha \in \{0.1, 1, 1.5, 5\}$. Plot $\left\|x^{(k)} - x^*\right\|_2$ on a logarithmic scale.

(c) Explain why $\alpha = 1$ reaches the answer immediately, why $0 < \alpha < 2$ converges here, and why $\alpha = 5$ diverges.

13. Gradient descent, also called steepest descent in the Euclidean norm, chooses

$$d^{(k)} = -\nabla f(x^{(k)}), \qquad x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}.$$

With a constant-step method, $\alpha_k = \alpha$ is fixed before the iterations. An unknown quadratic has gradient

$$\nabla f(x) = \begin{bmatrix} 6x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{bmatrix}.$$

Starting from $x^{(0)} = (3, -1)^{\mathsf{T}}$, compute two iterations with $\alpha = 1/4$ using only the gradients. Then test $\alpha = 0.02, 0.03, \ldots, 0.40$ for 40 iterations and plot the final gradient norm. Derive the stability interval and the best worst-case constant step from the Hessian eigenvalues.

14. Exact line search gradient descent uses the same direction $d^{(k)} = -\nabla f(x^{(k)})$, but computes a new optimal step at every iteration:

$$\alpha_k \in \arg\min_{\alpha \geq 0} f\big(x^{(k)} + \alpha d^{(k)}\big).$$

That is, at each iteration, the scalar α is **optimally computed** by minimizing the function $f\big(x^{(k)} + \alpha d^{(k)}\big)$ which is only a function of α while the other quantities are actually constant. Thus constant step gradient descent chooses α in advance, whereas exact line search solves a one dimensional optimization problem during each iteration. For

$$f(x) = (x_1 - 2)^4 + 2(x_2 + 1)^2 + 3(x_3 - 3)^4, \qquad x^{(0)} = (2, 0, 5)^\mathsf{T},$$

perform three exact-line-search iterations. Then implement the method. Stop when the successive-step norm is at most 10^{-6}, report the iteration count, and plot both the objective value and step norm versus iteration.

Hint: define $\phi_k(\alpha) = f(x^{(k)} + \alpha d^{(k)})$. Exact line search can be computed by solving $\phi_k'(\alpha) = 0$ for a nonnegative minimizer, or numerically with `fminbnd` in MATLAB or `minimize_scalar` in Python.

15. The general LP formulation used in class is

$$\begin{array}{ll} \text{minimize}_x & \tilde{c}^\mathsf{T} x + c_0 \\ \text{subject to} & \hat{A}x \leq \hat{B}, \\ & \tilde{A}x \geq \tilde{b}, \\ & Ex = d, \\ & \ell_j \leq x_j \leq u_j \quad \text{when bounds exist}, \\ & x_j \text{ free for } j \in \mathcal{F}. \end{array}$$

(a) Derive from it the equality standard form

$$\begin{array}{ll} \text{minimize}_z & c^\mathsf{T} z + \kappa \\ \text{subject to} & Az = b, \\ & z \geq 0. \end{array}$$

Hints: use $x_j = x_j^+ - x_j^-$ for a free variable; use a shift such as $x_j = \ell_j + z_j$ for a lower bound; add a nonnegative slack to a $\leq$ constraint; and subtract a nonnegative surplus variable from a $\geq$ constraint.

(b) Convert the following LP completely and recover its optimizer in both representations:

$$\begin{array}{ll} \text{minimize} & 2x_1 + x_2 \\ \text{subject to} & x_1 + 2x_2 \geq 3, \quad -x_1 + x_2 \leq 4, \\ & x_1 \text{ free}, \quad -1 \leq x_2 \leq 3. \end{array}$$

16. A planning team is screening one regional Walmart store to serve six Upper Cumberland communities. The attached file is named `upper_cumberland_coordinates.csv` and is located in the `data` folder. It contains the following local coordinates in kilometers and population weights:

Community	East a_{i1}	North a_{i2}	Population
Algood	5.75	5.74	3963
Baxter	-11.11	0.60	1578
Monterey	22.10	-0.40	2746
Livingston	16.53	26.56	3905
Gainesboro	-12.42	23.69	920
Sparta	3.43	-23.70	4998

Let $s = (s_1, s_2)$ be the store location and w_i be community i's population divided by total population. The candidate parcel must satisfy $-8 \le s_1 \le 12$ and $-5 \le s_2 \le 15$.

For each community, introduce an epigraph variable t_i. The inequality $\|s - a_i\|_2 \le t_i$ makes t_i an upper bound on that community's distance; because t_i is minimized, it equals the distance at an optimum. Introduce r with $t_i \le r$ for every i, so r is the largest distance.

(a) Write the complete SOCP that minimizes $\sum_i w_i t_i + 0.15r$.

(b) Solve it and report the store location, every t_i, and r.

(c) Make a two-dimensional network plot: draw each town as a labeled node, scale node size by population, draw the permitted parcel rectangle, mark the optimal store with a star, and connect the store to every town with an edge.

(d) Explain what changing the fairness coefficient from 0.15 toward zero or toward one would do.

17. Five projects have binary selections $x \in \{0,1\}^5$, costs $a = (6,5,4,3,2)$, benefits $b = (9,8,6,5,7)$, and budget 10. Consider

$$\text{minimize } -b^\mathsf{T} x - 4x_1x_2 - 3x_3x_4 + 2x_2x_5 \quad \text{subject to} \quad a^\mathsf{T} x \le 10, quad x \in \{0,1\}^5.$$

For each model below, write a complete convex formulation and solve it in CVX/CVXPY:

(a) R1: replace each displayed product by y_{ij}, use $0 \le x \le 1$, and impose only $0 \le y_{ij} \le x_i, x_j$;

(b) R2: add the remaining McCormick inequality $y_{ij} \ge x_i + x_j - 1$;

(c) R3: add the valid cover cuts $\sum_{i \in C} x_i \le |C| - 1$ for

$$C \in \{\{1,2\}, \{1,3,5\}, \{1,4,5\}, \{2,3,4\}, \{2,3,5\}\};$$

(d) R4: introduce $X \approx xx^\mathsf{T}$, impose $\begin{bmatrix} X & x \\ x^\mathsf{T} & 1 \end{bmatrix} \succeq 0$, $\operatorname{diag}(X) = x$, all pairwise McCormick inequalities, the cover cuts, and both RLT products of $a^\mathsf{T} x \le 10$ with x_j and $1 - x_j$.

Enumerate the 32 binary vectors to find the true optimum. Make a table of all four lower bounds and the true value. Also submit a two-panel figure: a bar chart comparing the four bounds with the true optimum, and a heat map showing the five components of the relaxed x from R1–R4 beside the exact binary solution. Explain why a tighter lower bound is useful.

18. *Ridge regression* is linear least squares regression with an ℓ_2 penalty that discourages large coefficients. Let a_i^T be the standardized feature row, including a leading 1 for the intercept, and let y_i be demand. The ridge problem is

$$\underset{\theta}{\text{minimize}} \ \frac{1}{2N} \sum_{i=1}^{N} (a_i^\mathsf{T}\theta - y_i)^2 + \frac{\lambda}{2} \|D\theta\|_2^2, \qquad D = \operatorname{diag}(0,1,\ldots,1).$$

The zero in D leaves the intercept unpenalized. Use the attached file `SeoulBikeData.csv` in the `data` folder. Keep functioning days. Predict hourly rented-bike count from temperature, humidity, wind speed, visibility, solar radiation, rainfall, snowfall, $\sin(2\pi h/24)$, $\cos(2\pi h/24)$, and a holiday indicator.

(a) Show that the ridge objective is convex and derive its gradient and normal equations.

(b) Use the first 70% of rows for training, the next 15% for validation, and the final 15% for testing. Standardize using training statistics only. Choose $\lambda \in \{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$ by validation RMSE.

(c) Report test RMSE and MAE, compare them with the training/validation-mean baseline, and plot the first 168 test predictions against the observations.

(d) Discuss temporal leakage and at least two limitations of this linear model.

19. *Logistic regression* models a binary response through

$$p_i = \Pr(y_i = 1 \mid a_i) = \sigma(a_i^\mathsf{T}\theta), \qquad \sigma(s) = \frac{1}{1+e^{-s}}.$$

The ℓ_2-regularized logistic loss is

$$\underset{\theta}{\text{minimize}} \ \frac{1}{N}\sum_{i=1}^{N}\left[\log\left(1+e^{a_i^\mathsf{T}\theta}\right) - y_i a_i^\mathsf{T}\theta\right] + \frac{\lambda}{2}\left\|D\theta\right\|_2^2, \qquad D = \operatorname{diag}(0,1,\ldots,1).$$

The first entry of θ is the unpenalized intercept. Use the attached files `datatraining.txt` and `datatest.txt` in the `data` folder. The response is room occupancy; predictors are temperature, humidity, light, CO_2, and humidity ratio.

(a) Derive the gradient and Hessian of the displayed loss and explain why the training problem is convex.

(b) Standardize using only the training file, set $\lambda = 0.01$, and fit the model with CVX/CVXPY, Newton's method, or L-BFGS (https://en.wikipedia.org/wiki/Limited-memory_BFGS).

(c) At probability threshold 0.5, report the confusion matrix, accuracy, precision, and recall. Plot the confusion matrix with readable class labels and counts.

(d) Explain why accuracy alone can be misleading and name one concern when the model is used for HVAC control.