Instructions: Solutions are due on Monday, February 10th, 2014, 11:59:59 EDT. All your solutions should be prepared in \LaTeX{} and the PDF and .tex should be submitted to Thibaut. For each question, the best and correct answers will be selected as sample solutions for the entire class to enjoy. If you prefer that we do not use your solutions, please indicate this clearly on the first page of your assignment.

1. Convex Sets. Prove or give a counterexample:
   a. The intersection of convex sets is a convex set.
   b. Every polyhedron is a convex set.

2. Infeasibility and Unboundedness. Discuss the feasibility and boundedness of the following linear programs:

   maximize $2x_2 + x_3$  \hspace{1cm} minimize $x + y + z + w$

   subject to $x_1 - x_2 \leq 5$ \hspace{2cm} subject to $x + 3y + 2z + 4w \leq 5$

   $-2x_1 + x_2 \leq 3$ \hspace{2cm} $3x + y + 2z + w \leq 4$

   $x_1 - 2x_3 \leq 5$ \hspace{2cm} $5x + 3y + 3z + 3w = 9$

   $x_1, x_2, x_3 \geq 0$ \hspace{2cm} $x, y, z, w \geq 0$

3. Projection onto Convex Sets. In this exercise we will prove that for every closed convex set $C \subseteq \mathbb{R}^n$ and every point $y \in \mathbb{R}^n$ there is a unique point in $C$ closest to $y$, the projection of $y$ onto $C$. We will also give a geometric characterization of this point.

   a. Prove that a strictly convex function has at most one global minimum.

   b. Prove that the distance function $f_y(x) = (\|y - x\|_2)^2$ is continuous and strictly convex. Recall that $f$ is strictly convex over the convex set $S$ iff for any two points $x, y \in S$ with $x \neq y$ and $\lambda \in (0, 1)$, $f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$.

   c. Prove that for every nonempty closed convex set $C \subseteq \mathbb{R}^n$ and $y \in \mathbb{R}^n$ there exists a unique point $z \in C$ s.t. $z = \arg\min_{x \in C} \|y - x\|_2^2$. (Hint: show that you can restrict yourself to a closed bounded subset of $C$).

   d. Does the above statement hold when $C$ is not closed? Prove or give a counterexample.

   e. For $C$ and $y \in \mathbb{R}^n$ as above, prove that $z = \arg\min_{x \in C} \|y - x\|_2^2$ if and only if $(y - z)^T(x - z) \leq 0$, $\forall x \in C$. Draw a picture illustrating this property. (Hint: Recall the law of the parallelogram: $\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2$)
4. Separating Hyperplanes. In this exercise we will show that for every nonempty, closed, convex set $C$ and $y \notin C$, there exists a hyperplane separating $C$ and $y$.

   a. Prove that for every non-empty, closed, convex set $C$ and $y \notin C$ there exists a point $a \in \mathbb{R}^n$, $a \neq 0$ and $\alpha \in \mathbb{R}$ s.t. $a^T x < \alpha < a^T y$. Hint: use $a = (y - z)$ for $z = \text{argmin}_{x \in C} \left( \|y - x\| \right)^2$.

   b. Does the above theorem hold for sets that are not convex? Prove or give a counterexample.

   c. Given the set $C = \{(x_i, x_j) \in \mathbb{R}^2 : 3 \leq x_i, x_j \leq 5\}$ and $y = (1, 1)$, find a separating hyperplane between $C$ and $y$.

5. Multivariate Linear Regression. In this exercise, we will study an extension of the single-dimensional linear regression model (seen in class) to the multivariate case. Given $n$ points $(\alpha_i, \beta_i) \in \mathbb{R}^d \times \mathbb{R}$ for $i \in \{1, \ldots, n\}$, we assume that $\beta_i$ can be expressed as a linear combination of $\alpha_i$’s entries:

   $\forall i \in \{1, \ldots, n\}, \beta_i = \alpha_i^T x$

for an unknown $x \in \mathbb{R}^d$. Note that the main difference with the model seen in class is that $\alpha_i$ and $x$ now belong to a $d$-dimensional space.

   In this $d$-dimensional model, the residual sum of squares takes the form $RSS(x) := \sum_{i=1}^{n} (\beta_i - \alpha_i^T x)^2$, and the optimization problem again consists in finding $x^* := \text{arg min}_x RSS(x)$.

   a. Show that $RSS$ can be written more concisely in matrix form: $RSS(x) = \|B - Ax\|_2^2$, for some vector $B \in \mathbb{R}^n$ and matrix $A \in \mathbb{R}^{n \times d}$ that you will define in terms of $(\alpha_i, \beta_i)$, $i \in \{1, \ldots, n\}$.

   b. Compute the gradient and the Hessian matrix of $RSS$. Hint: expand $RSS(x + h)$ for given $x$ and $h$.

   c. Show that $RSS$ is a convex function.

   d. Show that any critical point of $RSS$ is a local minimum. Recall that a critical point of a function $f$ is a vector $x$ such that $\nabla f(x) = 0$, where $\nabla f$ denotes the gradient of $f$.

   e. Under which (necessary and sufficient) condition on $A$ does $RSS$ admit a unique global minimum? How would you verify this condition algorithmically? Give the expression of the unique global minimum when this condition is satisfied.

6. Salary Prediction. In this exercise, the goal is to apply regression techniques for predictions.

   **Code:** You will need to write a few scripts (computer programs) to conduct your analysis and to create visual plots. You may use any programming language you like. Please attach the source code of your programs (clearly labeled) to your homework solutions.

   a. Download the dataset at [http://thibaut.horel.org/salary.dat](http://thibaut.horel.org/salary.dat). The format is tab-separated, the first line indicating the titles of the column (sx=sex, rk=rank, yr=number of years in current rank, dg=highest degree, yd=number of years since highest degree, sl=yearly salary).

   b. Plot the yearly salary as a function of the number of years in current rank.

   c. Compute the regression line of the previous plot. Report the parameters of its equation and plot it with the original data.

   d. Compute the RMSE (Root Mean Square Error): $RMSE := \sqrt{\frac{1}{n} \sum_{i=1}^{n} (s_i - \hat{s}_i)^2}$, where $s_i$ are the true salaries, $\hat{s}_i$ are the salaries predicted by your model, and $n$ is the number of data points.

   **Bonus.** Improve the quality of the prediction by considering more explanatory variables.