Homework 7
(due 02:00, Tuesday, December 4th, 2018)

Since the final midterm is comprehensive, I've added questions from the beginning of the semester.

1. **Orthogonal bases**

    1. **Gram-Schmidt orthogonalization:** Does the Gram-Schmidt orthogonalization process result in a set of basis vectors $\{g_n\}_{n=1}^{N}$ where each $g_n \in \mathbb{Q}^N$(the set of rationals in an N dimensional space) if one begins with a set of vectors $\{w_n\}_{n=1}^{N}$ with each $w_n \in \mathbb{Q}^N$? Give a brief conceptual justification.

    2. If not, modify the Gram-Schmidt orthogonalization procedure to produce a set of basis vectors $\{g_n\}_{n=1}^{N}$ where each $g_n \in \mathbb{Q}^N$. Implement the procedure on the Hendrix image (considered as a matrix) uploaded on to Canvas and used in Homework 3 (hendrix_final.png).

    3. Consider a vector $x \in \mathbb{R}^N$. (a) Show that there exist coefficients $\{b_n\}_{n=1}^{N}$ with each $b_n \in \mathbb{Q}$ and an orthogonal basis $\{g_n\}_{n=1}^{N}$ where each g_n is in $\mathbb{Q}^N$ and an $\epsilon > 0$ such that $\|x - \sum_{n=1}^{N} b_n g_n\| < \epsilon$. (b) Why is this useful?

2. **Linear Assignment problem:**

    1. Compare the results obtained from your own implementation of the Bregman divergence approach to linear assignment against the state of the art Jonker-Volgenant implementation (LAPJV or equivalent) in terms of speed and accuracy. Execute both algorithms on 10 randomly generated 100×100 assignment benefit matrices A. Document the parameter settings: (i) α and the convergence criteria for (ii) Sinkhorn balancing and (iii) overall convergence. [Make sure to report the final value of the (negative) objective $\sum_{ia} A_{ia} P_{ia}$ where P is a permutation obtained from both approaches.]

    2. Derive a relationship between the Lagrange parameters μ, ν and the Sinkhorn balancing procedure. Your derivation should be of the form $\mu^{(k+1)} = \mu^{(k)} + F^{(k)}(A, V, W, \mu, \nu)$ with a similar relationship for ν.

3. Is the Bregman divergence $\phi(y) - \phi(x) - (y - x)\phi'(x)$ convex w.r.t. both x and y [for a three times differentiable $\phi(x)$] ? Mathematically justify your answer.