HW 1

Due: Monday, Sep 9.

Remember that you may (and should!) talk about the problems amongst yourselves, or discuss them with me, providing attribution for any good ideas you might get – but your final write-up should be your own.

1: A problem of performance. In addition to dense matrices, MATLAB supports a sparse matrix data structure, in which only the nonzero elements of the matrix are stored. For a variety of square matrix size n and sparsity levels s (where s is the fraction of the entries that are nonzero), compare the speed of dense matrix-vector multiply and sparse matrix-vector multiply. You can use `As = sparse(A)` to make a sparse version of a dense matrix A. What do you observe about the relative performance of these options?

Note: Your performance will vary depending on your machine and your version of MATLAB, of course.

2: Identity plus low-rank. Suppose $u, v \in \mathbb{R}^n$ and $A = I + uv^T$. Write a function to compute $A^k x$ in $O(n)$ time. Your code should have the signature

```
function [Akx] = p2pow(u, v, k, x)
```

Note: If `c` is a scalar, you may assume `c^k` is computed in $O(1)$ time.

3: Structured perturbations. In this problem, we consider the structure of the groups of orthogonal and symplectic matrices:

1. Suppose $Q : \mathbb{R} \to \mathbb{R}^{n\times n}$ is a differentiable matrix-valued function and that $Q(t)$ is orthogonal for all t. Show that $S(t) = Q(t)^T \dot{Q}(t)$ is a skew-symmetric matrix, i.e. $S(t) = -S(t)^T$.

2. Suppose $S : \mathbb{R} \to \mathbb{R}^{2n\times 2n}$ is a differentiable matrix-valued function and that $S(t)$ is *symplectic* for all t, i.e.

$$S^T J S = J, \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

Show that $M = S^{-1}\dot{S}$ is *Hamiltonian*, i.e. $JM + M^T J = 0$.

4: Norm! Suppose $W \in \mathbb{R}^{m \times n}$ is *nonnegative* (all the entries are nonnegative) and *row stochastic* (each row sums to one), and let $p \geq 1$ be fixed. The function $\alpha \mapsto \alpha^p$ is *convex*, so if $y = Wx$, then

$$|y_i|^p \leq \sum_j w_{ij} |x_j|^p.$$

Argue that

1. If $u_j = \sum_i w_{ij}$ and $v_j = |x_j|^p$, then $\|y\|_p^p \leq u \cdot v$.
2. Therefore, $\|y\|_p^p \leq \|u\|_\infty \|v\|_1$.
3. Therefore, $\|y\|_p^p \leq \|W\|_1 \|x\|_p^p$.
4. Therefore, $\|W\|_p \leq \|W\|_1^{1/p}$.

Note:

1. Each of the steps above should be explained in a few words or symbols. Don't overthink!
2. The more general statement (which you do not need to prove) is that for any matrix A, $\|A\|_p \leq \|A\|_1^{1/p} \|A\|_\infty^{1/q}$, where $1/p + 1/q = 1$. If you're curious, you can see the whole proof on pages 28–29 of Kato's *Perturbation Theory of Linear Operators*.
3. This question was inspired by a query by Bobby Kleinberg about two years ago, which I take as evidence that (i) this fact is interesting to someone beside me and (ii) it is not obvious!