# PS 2

Due: Weds, Feb 4

**1: By the book** Book section 4.6, problems 5-10

**4.6.5** We can show by counterexample that $\|A^{-1}\|$ and $\|A\|^{-1}$ need not necessarily be equal. Consider the diagonal matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

We have $\|A\| = 2$ and $\|A^{-1}\| = 1 \neq 1/2$ in any of the operator $p$-norms.

**4.6.6** We prove $\|A\|_1$ is the maximum absolute column sum by an argument similar to the proof in the book (pages 76–77) that $\|A\|_\infty$ is the maximum absolute row sum. First, let $\mathbf{x}$ be any vector with $\|\mathbf{x}\|_1 = 1$, we have

$$\|A\mathbf{x}\|_1 = \sum_i \left| \sum_j a_{ij} x_j \right| \leq \sum_{i,j} |a_{ij}||x_j| \leq \max_j \sum_i |a_{ij}|,$$

where the last inequality follows from the fact that a weighted average of $n$ numbers is bounded by the largest of those numbers. Conversely, let $k$ denote the column of $A$ that maximizes $\sum_i |a_{ij}|$, and let $\mathbf{e}_k$ be column $k$ of the identity matrix (i.e. the vector that has a one in entry $k$ and zeros elsewhere). Then

$$\|A\|_1 \geq \|A\mathbf{e}_k\|_1 = \max_j \sum_i |a_{ij}|.$$

Combining the two inequalities, we have

$$\max_j \sum_i |a_{ij}| \leq \|A\mathbf{e}_k\|_1 \leq \|A\|_1 \leq \max_j \sum_i |a_{ij}|$$

and this implies $\|A\|_1 = \max_j \sum_i |a_{ij}|$, as desired.

**4.6.7** If $A$ is positive definite, we can write $A = S^T S$ for some nonsingular $S$. The simplest way to do this is via the eigenvalue decomposition ($A = Q\Lambda Q^T$, $S = \Lambda^{1/2} Q^T$); later, we will see the Cholesky decomposition also does the trick. Therefore,

$$\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A \mathbf{x}} = \sqrt{\mathbf{x}^T S^T S \mathbf{x}} = \|S\mathbf{x}\|_2$$

If $S$ is a nonsingular square matrix and $\mathbf{x} \mapsto \|\mathbf{x}\|$ is any norm, then $\mathbf{x} \mapsto \|S\mathbf{x}\|$ is also a norm. That means $\mathbf{x} \mapsto \|S\mathbf{x}\|$ inherits the three properties of the original norm:

1. Positive definiteness: that $\|S\mathbf{x}\| \geq 0$ for any $\mathbf{x}$ follows from $\|\cdot\|$ being a norm; that equality holds only when $\mathbf{x} = 0$ follows from non-singularity of $S$ (i.e. $S\mathbf{x} = 0$ iff $\mathbf{x} = 0$)
2. Homogeneity: $\|S(\alpha\mathbf{x})\| = \|\alpha(S\mathbf{x})\| = |\alpha|\|S\mathbf{x}\|$
3. Triangle inequality: $\|S(\mathbf{x}+\mathbf{y})\| = \|S\mathbf{x} + S\mathbf{y}\| \leq \|S\mathbf{x}\| + \|S\mathbf{y}\|$

**4.6.8**

1. Permutation Matrices
   We write elements of a permutation matrix $P$ as $p_{ij}$. For $P$ to be orthogonal we must have $P^T P = I$, or equivalently $\sum_j p_{ji}p_{jk} = \delta_{ik}$. This equality trivially follows since when $i = k$ we have terms given by $p_{ji}p_{ji}$ and for each $i$ there is one and only one $j$ such that this is equal to 1.
2. SPD/Nonsingular/Diagonal Matrices
   The matrix
   $$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$
   is diagonal, nonsingular, and SPD, but is not orthogonal.

**4.6.9** For the matrix

$$A = A(a,b) = \begin{bmatrix} a & 1 & 1+b \\ 1 & a & 1 \\ 1-b^2 & 1 & a \end{bmatrix}$$

to be SPD we must have (1) symmetry ($A = A^T$) and (2) positive definiteness ($\mathbf{x}^\mathbf{T} A\mathbf{x} > 0$). For symmetry, we require that $a_{13} = a_{31}$, which implies $b$ must equal 0 or $-1$. To determine the values of $a$ so that (2) holds, we observe that the eigenvalues of $A(a,b)$ are the eigenvalues of $A(0,b)$ plus $a$. One of the equivalent conditions for positive definiteness is that all eigenvalues are positive, so we require $a > \lambda_{\min}(A(0,b))$. For $b = 0$ this gives $a > 1$ and for $b = -1$ we have $a > \sqrt{2}$.

**4.6.10**

1. Orthogonality

$$C^T C = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c^2+s^2 & 0 \\ 0 & c^2+s^2 \end{bmatrix}.$$

   Thus, we require $c^2 + s^2 = 1$ for $C$ to be orthogonal.

2. Givens rotations
   We have the equation $C\mathbf{a} = \mathbf{a}' = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$. Orthogonality implies we must have $\|\mathbf{a}\| = \|\mathbf{a}'\| = |\alpha|$. Thus, $\alpha^2 = {a_1}^2 + {a_2}^2$. To find expressions for $s$ and $c$ we multiply through to get our additional equations of constraint. Solving $a_1 c + a_2 s = \alpha$, $-a_1 s + a_2 c = 0$, $s^2 + c^2 = 1$, and $\alpha^2 = {a_1}^2 + {a_2}^2$ for $c$ and $s$ gives $s = \frac{a_2}{\alpha}$ and $c = \frac{a_1}{\alpha}$.

**2: Recognizing Rank**

1. For $n = 3$,

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

2. For a general $n \times n$ matrix $A$, the entries are $a_{ij} = i + (j-1)n$. Let $\mathbf{e}$ be the vector of all ones, and let $[n]$ denote the vector with entries 1 through $n$; then

$$A = [n]\mathbf{e}^T + \mathbf{e}n([n] - \mathbf{e})^T$$

   Thus, $A$ is a sum of two independent rank one contributions (i.e. $A$ is rank two).

3. To rewrite `ps2mult()` so that it runs in $O(n)$ time we use the low-rank decomposition described above. The resulting algorithm requires roughly $2n$ multiplies and $3n$ additions. A performance comparison plot and code are shown below.

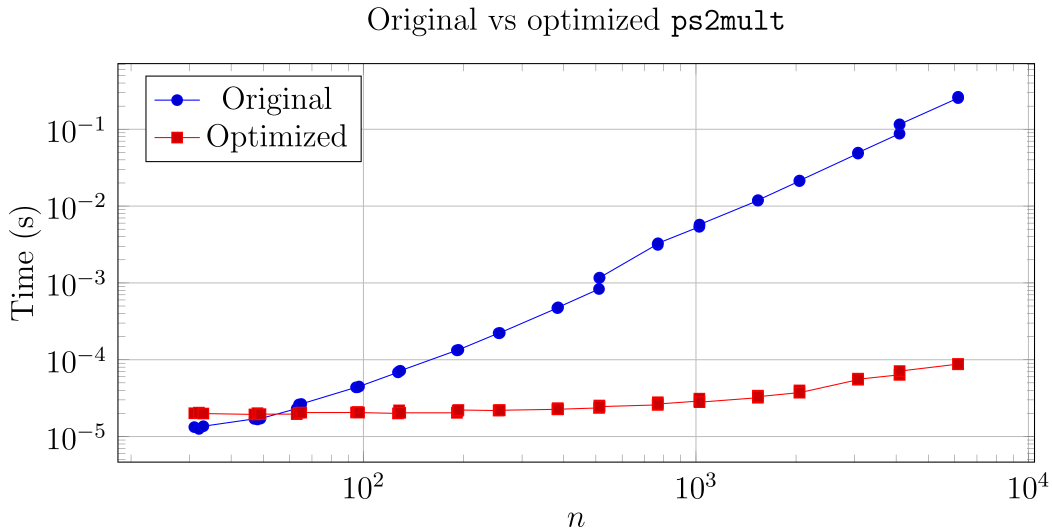


Figure 1: Timing of original ($O(n^2)$) and optimized ($O(n)$) codes. The original code is clearly quadratic (slope 2 on a log-log plot); the optimized code is fast enough that the time is dominated by constant overheads, resulting in a nearly flat curve for $n < 1000$.

```
function [y] = ps2mult(x)

  % a(i,j) = i + (j-1)*n
  % A = idx * e' + n * e * (idx-1)
  % where idx = 1:n

  n = length(x);
  y = (1:n)' * sum(x);
  y = y + n * ((0:n-1) * x);
```