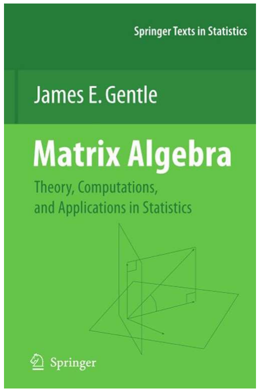
Theory of Matrices

Chapter 5. Matrix Transformations and Factorizations

5.3. Householder Transformations (Reflections)—Proofs of Theorems



Theorem 5.3.1

Theorem 5.3.1. For nonzero $x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$, define

$$q = [x_1, x_2, \ldots, x_{k-1}, x_k + \mathrm{sgn}(x_k)\|x\|_2, x_{k+1}, \ldots, x_n]^T$$

and let $u = q/\|q\|_2$. Then $H = I - 2uu^T$ maps x to $Hx = \tilde{x}$ where all entries of $\tilde{x}$ are 0 except for the kth entry (when $x_k \neq 0$).

Proof. First, notice that

$$\|q\|_2^2 = x_1^2 + x_2^2 + \cdots + x_{k-1}^2 + (x_k^2 + 2x_k\mathrm{sgn}(x_k)\|x\|_2 + \mathrm{sgn}(x_k)^2\|x\|_2^2)$$
$$+x_{k+1}^2 + \cdots x_n^2 = \|x\|_2^2 + 2|x_k|\|x\|_2 + \|x\|_2^2 = 2\|x\|_2(\|x\|_2 + |x_k|).$$

Notice that the (i,j) entry of qq^T is

$$q_iq_j = \begin{cases} x_ix_j & \text{for } i \neq k \neq j \\ (x_k + \mathrm{sgn}(x_k)\|x\|_2)x_j & \text{for } i = k, j \neq k \\ x_i(x_k + \mathrm{sgn}(x_k)\|x\|_2) & \text{for } i \neq k, j = k \\ (x_k + \mathrm{sgn}(x_k)\|x\|_2)^2 & \text{for } i = j = k. \end{cases}$$

Let $y_k = x_k + \mathrm{sgn}(x_k)\|x\|_2$.

Theorem 5.3.1 (continued 1)

Proof (continued). Now $qq^T =$

$$\begin{bmatrix} x_1x_1 & x_1x_2 & \cdots & x_1x_{k-1} & x_1y_k & x_1x_{k+1} & \cdots & x_1x_n \\ x_2x_1 & x_2x_2 & \cdots & x_2x_{k-1} & x_2y_k & x_2x_{k+1} & \cdots & x_2x_n \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ x_{k-1}x_1 & x_{k-1}x_2 & \cdots & x_{k-1}x_{k-1} & x_{k-1}y_k & x_{k-1}x_{k+1} & \cdots & x_{k-1}x_n \\ y_kx_1 & y_kx_2 & \cdots & y_kx_{k-1} & y_ky_k & y_kx_{k+1} & \cdots & y_kx_n \\ x_{k+1}x_1 & x_{k+1}x_2 & \cdots & x_{k+1}x_{k-1} & x_{k+1}y_k & x_{k+1}x_{k+1} & \cdots & x_{k+1}x_n \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & x_nx_{k-1} & x_ny_k & x_nx_{k+1} & \cdots & x_nx_n \end{bmatrix}$$

...

Theorem 5.3.1 (continued 2)

Proof (continued). ...and $qq^Tx =$

$$\begin{bmatrix} x_1\sum_{\ell=1}^n x_\ell^2 - x_1x_k^2 + x_1y_kx_k \\ x_2\sum_{\ell=1}^n x_\ell^2 - x_2x_k^2 + x_2y_kx_k \\ \vdots \\ x_{k-1}\sum_{\ell=1}^n x_\ell^2 - x_{k-1}x_k^2 + x_{k-1}y_kx_k \\ y_k\sum_{\ell=1}^n x_\ell^2 - y_kx_k^2 + y_k^2x_k \\ x_{k+1}\sum_{\ell=1}^n x_\ell^2 - x_{k+1}x_k^2 + x_{k+1}y_kx_k \\ \vdots \\ x_n\sum_{\ell=1}^n x_\ell^2 - x_nx_k^2 + x_ny_kx_k \end{bmatrix} = \begin{bmatrix} x_1\|x\|_2^2 - x_1x_k(x_k - y_k) \\ x_2\|x\|_2^2 - x_2x_k(x_k - y_k) \\ \vdots \\ x_{k-1}\|x\|_2^2 - x_{k-1}x_k(x_k - y_k) \\ y_k\|x\|_2^2 - y_kx_k(x_k - y_k) \\ x_{k+1}\|x\|_2^2 - x_{k+1}x_k(x_k - y_k) \\ \vdots \\ x_n\|x\|_2^2 - x_nx_k(x_k - y_k) \end{bmatrix}$$

...

Theorem 5.3.1 (continued 3)

Proof (continued). . . . and since $y_k = x_k + \text{sgn}(x_k)\|x\|_2$ then $x_k(x_k - y_k) = x_k(-\text{sgn}(x_k)\|x\|_2) = -|x_k|\|x\|_2$ and

$$qq^T x = \begin{bmatrix} x_1\|x\|_2^2 + x_1|x_k|\|x\|_2 \\ x_2\|x\|_2^2 + x_2|x_k|\|x\|_2 \\ \vdots \\ x_{k-1}\|x\|_2^2 + x_{k-1}|x_k|\|x\|_2 \\ y_k\|x\|_2^2 + y_k|x_k|\|x\|_2 \\ x_{k+1}\|x\|_2^2 + x_{k+1}|x_k|\|x\|_2 \\ \vdots \\ x_n\|x\|_2^2 + x_n|x_k|\|x\|_2 \end{bmatrix}.$$

Theorem 5.3.1 (continued 4)

Proof (continued). Since $\|q\|_2^2 = 2\|x\|_2(\|x\|_2 + |x_k|)$ then $(I - 2uu^T)x = (I - 2qq^T/\|q\|_2^2)x = (\|q\|_2^2 I - 2qq^T)x/\|q\|_2^2 =$

$$\frac{1}{\|q\|_2^2}\begin{bmatrix} 2\|x\|_2(\|x\|_2 + |x_k|)x_1 - 2x_1\|x\|_2^2 - 2x_1|x_k|\|x\|_2 \\ 2\|x\|_2(\|x\|_2 + |x_k|)x_2 - 2x_2\|x\|_2^2 - 2x_2|x_k|\|x\|_2 \\ \vdots \\ 2\|x\|_2(\|x\|_2 + |x_k|)x_{k-1} - 2x_{k-1}\|x\|_2^2 - 2x_{k-1}|x_k|\|x\|_2 \\ 2\|x\|_2(\|x\|_2 + |x_k|)x_k - 2y_k\|x\|_2^2 - 2y_k|x_k|\|x\|_2 \\ 2\|x\|_2(\|x\|_2 + |x_k|)x_{k+1} - 2x_{k+1}\|x\|_2^2 - 2x_{k+1}|x_k|\|x\|_2 \\ \vdots \\ 2\|x\|_2(\|x\|_2 + |x_k|)x_n - 2x_n\|x\|_2^2 - 2x_n|x_k|\|x\|_2 \end{bmatrix} = \cdots$$

Theorem 5.3.1 (continued 5)

Proof (continued). . . .

$$= [0, 0, \ldots, 0, 2\|x\|_2(\|x\|_2+|x_k|)x_k - 2y_k\|x\|_2^2 - 2y_k|x_k|\|x\|_2, 0, \ldots, 0]^T/\|q\|_2^2$$

$$= [0, 0, \ldots, 0, 2\|x\|_2(\|x\|_2+|x_k|)x_k - 2\|x\|_2 y_k(\|x\|_2+|x_k|), 0, \ldots, 0]^T/\|q\|_2^2$$

$$\begin{aligned} &= [0, 0, \ldots, 0, 2\|x\|_2(\|x\|_2 + |x_k|)(x_k - y_k), 0, \ldots, 0]^T/\|q\|_2^2 \\ &= [0, 0, \ldots, 0, 2\|x\|_2(\|x\|_2 + |x_k|)(-\text{sgn}(x_k)\|x\|_2), 0, \ldots, 0]^T/\|q\|_2^2 \\ &\quad \text{since } x_k - y_k = -\text{sgn}(x_k)\|x\|_2 \\ &= \frac{[0, 0, \ldots, 0, 2\|x\|_2(\|x\|_2 + |x_k|)(-\text{sgn}(x_k)\|x\|_2), 0, \ldots, 0]^T}{2\|x\|_2(\|x\|_2 + |x_k|)} \\ &\quad \text{since } \|q\|_2^2 = 2\|x\|_2(\|x\|_2 + |x_k|) \\ &= [0, 0, \ldots, 0, -\text{sgn}(x_k)\|x\|_2, 0, \ldots, 0]^T. \end{aligned}$$

So the result holds. □