Problem 1 (Chapter 9, Exercise 2). Suppose $A$ is skew-Hermitian, i.e., $A^H = -A$. Prove that all its eigenvalues are pure imaginary.

Solution. Let $B = iA$, then $B^H = -iA^H = iA = B$ and therefore $B$ is Hermitian. Since $B$ has all real eigenvalues $\{\lambda_1, \cdots, \lambda_n\}$, all the eigenvalues of $A$ are of the form $\{-i\lambda_1, \cdots, -i\lambda_n\}$, and thus all pure imaginary.

Problem 2 (Chapter 9, Exercise 3). If $A$ is an $n \times n$ Hermitian matrix with eigenvalue $\lambda$ and corresponding right eigenvector $x$, show that $x$ is also a left eigenvector for $\lambda$. Prove the same result if $A$ is skew-Hermitian.

Solution. Let $x$ be a right eigenvector of $A$ with eigenvalue $\lambda$, then by definition

$$Ax = \lambda x$$

Taking Hermitian transpose on both sides,

$$x^HA^H = \bar{\lambda}x^H$$

- If $A$ is Hermitian, then $A^H = A$, and all eigenvalues are real, $\bar{\lambda} = \lambda$, so

$$x^HA = \lambda x^H$$

i.e. $x$ is also a left eigenvector of $A$ with the same eigenvalue $\lambda$.

- If $A$ is skew-Hermitian, then $A^H = -A$, and by Problem 9.2, all eigenvalues are purely imaginary, $\bar{\lambda} = -\lambda$, so

$$-x^HA = -\lambda x^H$$

i.e. $x$ is also a left eigenvector of $A$ with the same eigenvalue $\lambda$.

Problem 3. Suppose $A$ is positive semidefinite. Can you find a square root of this matrix? In other words, can you find a matrix $B$ such that $B^2 = A$? If yes, explain how you would construct it. If no, explain why no such matrix exists.

Solution. Yes, we can. Notice that $A$ can be written as $A = V \Sigma V^T$ where $V$ is unitary, $\Sigma$ is diagonal. The diagonal elements of $\Sigma$ are all the eigenvalues of $A$, and therefore are all non-negative. Consider $B = V D V^T$ where $D$ is a diagonal matrix with diagonal elements $D_{ii} = \sqrt{\Sigma_{ii}}$. Then we have $B^2 = V D V^T V D V^T = A$. 
Problem 4. Suppose $A$ is positive semidefinite. Show that the maximum eigenvalue of $A$, denoted by $\lambda_{\text{max}}$, is given by the so-called Rayleigh quotient

$$\sup_{w \neq 0} \frac{w^\top A w}{w^\top w}$$

Solution. Let $w_0$ be an eigenvector of $\lambda_{\text{max}}$, then we have

$$\frac{w_0^\top A w_0}{w_0^\top w_0} = \lambda_{\text{max}}.$$

On the other hand, denote the spectral representation of $A$ by

$$\sum_{i=1}^n \lambda_i v_i v_i^\top,$$

then we have

$$\frac{w^\top A w}{w^\top w} = \frac{\sum_{i=1}^n \lambda_i \langle w, v_i \rangle^2}{\|w\|^2} \leq \frac{\lambda_{\text{max}} \sum_{i=1}^n \|\langle w, v_i \rangle\|^2}{\|w\|^2} = \lambda_{\text{max}}.$$

Therefore, we have

$$\sup_{w \neq 0} \frac{w^\top A w}{w^\top w} = \lambda_{\text{max}},$$

as desired.

Problem 5. (a) Consider an $n \times n$ positive definite matrix $A$ with a largest eigenvalue greater than the second largest. Consider the following algorithm: start with a random vector $x^{(0)} \in \mathbb{R}^n$ such that $\|x^{(0)}\|_2 = 1$ for $t = 1, 2, 3, \cdots$ recursively define:

$$x^{(t)} = Ax^{(t-1)}/\|Ax^{(t-1)}\|_2.$$

What does this algorithm converge to?

Solution. We claim that the algorithm will ‘almost’ converge to the eigenvector with respect to the largest eigenvalue.

Consider the spectral representation of $A = \sum_{i=1}^n \lambda_i v_i v_i^\top$ where $\{v_1, \cdots, v_n\}$ is an orthonormal basis of $A$, $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$ are all the eigenvalues. Then we have

$$Ax = \sum_{i=1}^n \lambda_i \langle x, v_i \rangle v_i.$$

Therefore, by applying $A$ and normalizing $n$ times, we have

$$x^{(k)} = \frac{1}{C_k} \sum_{i=1}^n \lambda_i^k \langle x^{(0)}, v_i \rangle v_i,$$

where $C$ is the normalizing constant.
Suppose $\langle x^{(0)} , v_1 \rangle \neq 0$ (which means the initial vector is not orthogonal to $v_1$), we claim that

$$\frac{\lambda_i^k \langle x^{(0)} , v_1 \rangle}{C_n} \to 1.$$  

Let $c_{ik} = \lambda_i^k \langle x^{(0)} , v_i \rangle$, then we have $C_k = \sqrt{\sum_{i=1}^{n} c_{ik}^2}$. Notice that

$$\frac{c_{1k}}{c_{ik}} = \frac{\lambda_1^k \lambda_i}{\lambda_i} \frac{\langle x^{(0)} , v_1 \rangle}{\langle x^{(0)} , v_i \rangle} \to \infty$$

for all $i > 1$, as $\lambda_1$ is the unique largest eigenvector.

Therefore, as $k \to \infty$, if $\langle x^{(0)} , v_1 \rangle > 0$, $x^{(k)} \to v_1$, if $\langle x^{(0)} , v_1 \rangle < 0$, $x^{(k)} \to -v_1$. In either case, the algorithm converges to the eigenvector with respect to the largest eigenvalue.

(The edge case is $\langle x^{(0)} , v_1 \rangle = 0$, which has probability 0 if we choose a direction uniformly at random. In that case, the algorithm will converge to $v_k$ or $-v_k$, where $k$ is the smallest integer such that $x$ is not orthogonal with $v_k$.)