3.5 [RV73, page 22] *Running average of a convex function.* Suppose $f : \mathbf{R} \to \mathbf{R}$ is convex, with $\mathbf{R}_+ \subseteq \mathbf{dom}\, f$. Show that its *running average* F, defined as

$$F(x) = \frac{1}{x}\int_0^x f(t)\, dt, \qquad \mathbf{dom}\, F = \mathbf{R}_{++},$$

is convex. You can assume f is differentiable.

Solution. F is differentiable with

$$\begin{aligned} F'(x) &= -(1/x^2)\int_0^x f(t)\, dt + f(x)/x \\ F''(x) &= (2/x^3)\int_0^x f(t)\, dt - 2f(x)/x^2 + f'(x)/x \\ &= (2/x^3)\int_0^x (f(t) - f(x) - f'(x)(t-x))\, dt. \end{aligned}$$

Convexity now follows from the fact that

$$f(t) \geq f(x) + f'(x)(t-x)$$

for all $x, t \in \mathbf{dom}\, f$, which implies $F''(x) \geq 0$.

3.8 *Second-order condition for convexity.* Prove that a twice differentiable function f is convex if and only if its domain is convex and $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbf{dom}\, f$. *Hint.* First consider the case $f : \mathbf{R} \to \mathbf{R}$. You can use the first-order condition for convexity (which was proved on page 70).

Solution. We first assume $n = 1$. Suppose $f : \mathbf{R} \to \mathbf{R}$ is convex. Let $x, y \in \mathbf{dom}\, f$ with $y > x$. By the first-order condition,

$$f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x).$$

Subtracting the righthand side from the lefthand side and dividing by $(y-x)^2$ gives

$$\frac{f'(y) - f'(x)}{y - x} \geq 0.$$

Taking the limit for $y \to x$ yields $f''(x) \geq 0$.

Conversely, suppose $f''(z) \geq 0$ for all $z \in \mathbf{dom}\, f$. Consider two arbitrary points $x, y \in \mathbf{dom}\, f$ with $x < y$. We have

$$\begin{aligned} 0 &\leq \int_x^y f''(z)(y-z)\, dz \\ &= \left(f'(z)(y-z)\right)\Big|_{z=x}^{z=y} + \int_x^y f'(z)\, dz \\ &= -f'(x)(y-x) + f(y) - f(x), \end{aligned}$$

i.e., $f(y) \geq f(x) + f'(x)(y-x)$. This shows that f is convex.

To generalize to $n > 1$, we note that a function is convex if and only if it is convex on all lines, *i.e.*, the function $g(t) = f(x_0 + tv)$ is convex in t for all $x_0 \in \mathbf{dom}\, f$ and all v. Therefore f is convex if and only if

$$g''(t) = v^T \nabla^2 f(x_0 + tv) v \geq 0$$

for all $x_0 \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$, and t satisfying $x_0 + tv \in \mathbf{dom}\, f$. In other words it is necessary and sufficient that $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbf{dom}\, f$.

3.16 For each of the following functions determine whether it is convex, concave, quasiconvex, or quasiconcave.

(a) $f(x) = e^x - 1$ on $\mathbf{R}$.
Solution. Strictly convex, and therefore quasiconvex. Also quasiconcave but not concave.

(b) $f(x_1, x_2) = x_1 x_2$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

which is neither positive semidefinite nor negative semidefinite. Therefore, f is neither convex nor concave. It is quasiconcave, since its superlevel sets

$$\{(x_1, x_2) \in \mathbf{R}^2_{++} \mid x_1 x_2 \geq \alpha\}$$

are convex. It is not quasiconvex.

(c) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \frac{1}{x_1 x_2} \begin{bmatrix} 2/(x_1^2) & 1/(x_1 x_2) \\ 1/(x_1 x_2) & 2/x_2^2 \end{bmatrix} \succeq 0$$

Therefore, f is convex and quasiconvex. It is not quasiconcave or concave.

(d) $f(x_1, x_2) = x_1/x_2$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{bmatrix}$$

which is not positive or negative semidefinite. Therefore, f is not convex or concave. It is quasiconvex and quasiconcave (*i.e.*, quasilinear), since the sublevel and superlevel sets are halfspaces.

(e) $f(x_1, x_2) = x_1^2/x_2$ on $\mathbf{R} \times \mathbf{R}_{++}$.
Solution. f is convex, as mentioned on page 72. (See also figure 3.3). This is easily verified by working out the Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} 2/x_2 & -2x_1/x_2^2 \\ -2x_1/x_2^2 & 2x_1^2/x_2^3 \end{bmatrix} = (2/x_2) \begin{bmatrix} 1 \\ -2x_1/x_2 \end{bmatrix} \begin{bmatrix} 1 & -2x_1/x_2 \end{bmatrix} \succeq 0.$$

Therefore, f is convex and quasiconvex. It is not concave or quasiconcave (see the figure).

(f) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, on $\mathbf{R}^2_{++}$.
Solution. Concave and quasiconcave. The Hessian is

$$\begin{aligned} \nabla^2 f(x) &= \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^\alpha x_2^{-\alpha-1} \end{bmatrix} \\ &= \alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \begin{bmatrix} -1/x_1^2 & 1/x_1x_2 \\ 1/x_1x_2 & -1/x_2^2 \end{bmatrix} \preceq 0. \end{aligned}$$

f is not convex or quasiconvex.

3.18 Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following.

(a) $f(X) = \mathbf{tr}\left(X^{-1}\right)$ is convex on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

(b) $f(X) = (\det X)^{1/n}$ is concave on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

Solution. Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in \mathbf{S}^n$.

$$\begin{aligned}
g(t) &= \mathbf{tr}((Z + tV)^{-1}) \\
&= \mathbf{tr}\left(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}\right) \\
&= \mathbf{tr}\left(Z^{-1}Q(I + t\Lambda)^{-1}Q^T\right) \\
&= \mathbf{tr}\left(Q^T Z^{-1}Q(I + t\Lambda)^{-1}\right) \\
&= \sum_{i=1}^n (Q^T Z^{-1} Q)_{ii}(1 + t\lambda_i)^{-1},
\end{aligned}$$

where we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. In the last equality we express g as a positive weighted sum of convex functions $1/(1+t\lambda_i)$, hence it is convex. Second part is similar.

3.20 *Composition with an affine function.* Show that the following functions $f : \mathbf{R}^n \to \mathbf{R}$ are convex.

(a) $f(x) = \|Ax - b\|$, where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and $\|\cdot\|$ is a norm on $\mathbf{R}^m$.

Solution. f is the composition of a norm, which is convex, and an affine function.

(b) $f(x) = -\left(\det(A_0 + x_1A_1 + \cdots + x_nA_n)\right)^{1/m}$, on $\{x \mid A_0 + x_1A_1 + \cdots + x_nA_n \succ 0\}$, where $A_i \in \mathbf{S}^m$.

Solution. f is the composition of the convex function $h(X) = -(\det X)^{1/m}$ and an affine transformation, To see that h is convex on $\mathbf{S}^m_{++}$, we restrict h to a line and prove that $g(t) = -\det(Z + tV)^{1/m}$ is convex:

$$\begin{aligned} g(t) &= -(\det(Z+tV))^{1/m} \\ &= -(\det Z)^{1/m}(\det(I + tZ^{-1/2}VZ^{-1/2}))^{1/m} \\ &= -(\det Z)^{1/m}(\prod_{i=1}^{m}(1+t\lambda_i))^{1/m} \end{aligned}$$

where $\lambda_1, \ldots, \lambda_m$ denote the eigenvalues of $Z^{-1/2}VZ^{-1/2}$. We have expressed g as the product of a negative constant and the geometric mean of $1 + t\lambda_i$, $i = 1, \ldots, m$. Therefore g is convex.

(c) $f(X) = \mathbf{tr}\,(A_0 + x_1A_1 + \cdots + x_nA_n)^{-1}$, on $\{x \mid A_0 + x_1A_1 + \cdots + x_nA_n \succ 0\}$, where $A_i \in \mathbf{S}^m$. (Use the fact that $\mathbf{tr}(X^{-1})$ is convex on $\mathbf{S}^m_{++}$; see exercise 3.18.)

Solution. f is the composition of $\mathbf{tr}\,X^{-1}$ and an affine transformation

$$x \mapsto A_0 + x_1A_1 + \cdots + x_nA_n.$$

3.26 *More functions of eigenvalues.* Let $\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X)$ denote the eigenvalues of a matrix $X \in \mathbf{S}^n$. We have already seen several functions of the eigenvalues that are convex or concave functions of X.

- The maximum eigenvalue $\lambda_1(X)$ is convex (example 3.10). The minimum eigenvalue $\lambda_n(X)$ is concave.

- The sum of the eigenvalues (or trace), $\mathbf{tr}\, X = \lambda_1(X) + \cdots + \lambda_n(X)$, is linear.
- The sum of the inverses of the eigenvalues (or trace of the inverse), $\mathbf{tr}(X^{-1}) = \sum_{i=1}^n 1/\lambda_i(X)$, is convex on $\mathbf{S}_{++}^n$ (exercise 3.18).
- The geometric mean of the eigenvalues, $(\det X)^{1/n} = (\prod_{i=1}^n \lambda_i(X))^{1/n}$, and the logarithm of the product of the eigenvalues, $\log\det X = \sum_{i=1}^n \log\lambda_i(X)$, are concave on $X \in \mathbf{S}_{++}^n$ (exercise 3.18 and page 74).

In this problem we explore some more functions of eigenvalues, by exploiting variational characterizations.

(a) *Sum of k largest eigenvalues.* Show that $\sum_{i=1}^k \lambda_i(X)$ is convex on $\mathbf{S}^n$. *Hint.* [HJ85, page 191] Use the variational characterization

$$\sum_{i=1}^k \lambda_i(X) = \sup\{\mathbf{tr}(V^TXV) \mid V \in \mathbf{R}^{n\times k},\ V^TV = I\}.$$

Solution. The variational characterization shows that f is the pointwise supremum of a family of linear functions $\mathbf{tr}(V^TXV)$.

(b) *Geometric mean of k smallest eigenvalues.* Show that $(\prod_{i=n-k+1}^n \lambda_i(X))^{1/k}$ is concave on $\mathbf{S}_{++}^n$. *Hint.* [MO79, page 513] For $X \succ 0$, we have

$$\left(\prod_{i=n-k+1}^n \lambda_i(X)\right)^{1/k} = \frac{1}{k}\inf\{\mathbf{tr}(V^TXV) \mid V \in \mathbf{R}^{n\times k},\ \det V^TV = 1\}.$$

Solution. f is the pointwise infimum of a family of linear functions $\mathbf{tr}(V^TXV)$.

(c) *Log of product of k smallest eigenvalues.* Show that $\sum_{i=n-k+1}^n \log\lambda_i(X)$ is concave on $\mathbf{S}_{++}^n$. *Hint.* [MO79, page 513] For $X \succ 0$,

$$\prod_{i=n-k+1}^n \lambda_i(X) = \inf\left\{\prod_{i=1}^n (V^TXV)_{ii} \;\middle|\; V \in \mathbf{R}^{n\times k},\ V^TV = I\right\}.$$

Solution. f is the pointwise infimum of a family of concave functions $\log\prod_i (V^TXV)_{ii}$.

3.43 *First-order condition for quasiconvexity.* Prove the first-order condition for quasiconvexity given in §3.4.3: A differentiable function $f : \mathbf{R}^n \to \mathbf{R}$, with $\mathbf{dom}\, f$ convex, is quasiconvex if and only if for all $x, y \in \mathbf{dom}\, f$,

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0.$$

Hint. It suffices to prove the result for a function on $\mathbf{R}$; the general result follows by restriction to an arbitrary line.

Solution. First suppose f is a differentiable function on $\mathbf{R}$ and satisfies

$$f(y) \leq f(x) \Longrightarrow f'(x)(y - x) \leq 0. \tag{3.43.A}$$

Suppose $f(x_1) \geq f(x_2)$ where $x_1 \neq x_2$. We assume $x_2 > x_1$ (the other case can be handled similarly), and show that $f(z) \leq f(x_1)$ for $z \in [x_1, x_2]$. Suppose this is false, *i.e.*, there exists a $z \in [x_1, x_2]$ with $f(z) > f(x_1)$. Since f is differentiable, we can choose a z that also satisfies $f'(z) < 0$. By (3.43.A), however, $f(x_1) < f(z)$ implies $f'(z)(x_1 - z) \leq 0$, which contradicts $f'(z) < 0$.

To prove sufficiency, assume f is quasiconvex. Suppose $f(x) \geq f(y)$. By the definition of quasiconvexity $f(x + t(y - x)) \leq f(x)$ for $0 < t \leq 1$. Dividing both sides by t, and taking the limit for $t \to 0$, we obtain

$$\lim_{t \to 0} \frac{f(x + t(y - x)) - f(x)}{t} = f'(x)(y - x) \leq 0,$$

which proves (3.43.A).