Solutions to the Exercises of Section 2.8.

2.8.1. Let $\mathbf{z} = (z_1,\ldots,z_{k+1})^T \in S_1$, $\mathbf{w} = (w_1,\ldots,w_{k+1})^T \in S_1$, and $0 \le \beta \le 1$. Then, $\mathbf{x} = (z_1,\ldots,z_k)^T \in S$ and $\mathbf{y} = (w_1,\ldots,w_k)^T \in S$, and $f(\mathbf{x}) \le z_{k+1}$ and $f(\mathbf{y}) \le w_{k+1}$. Since S is convex, $\beta\mathbf{x}+(1-\beta)\mathbf{y} \in S$, and since f is convex, $f(\beta\mathbf{x}+(1-\beta)\mathbf{y}) \le \beta f(\mathbf{x})+(1-\beta)f(\mathbf{y}) \le \beta z_{k+1}+(1-\beta)w_{k+1}$. Hence, $\beta\mathbf{z}+(1-\beta)\mathbf{w} = (\beta z_1+(1-\beta)w_1,\ldots,\beta z_{k+1}+(1-\beta)w_{k+1})^T \in S_1$.

2.8.2. (a) Let $f(x)$ be a convex function defined on an interval I of the real line, and let x_0 be an interior point of I. Find $x_1 \in I$ and $x_2 \in I$ such that $x_1 < x_0 < x_2$. Let $x_0 < x < x_2$. From convexity, we have

$$f(x) \le \frac{x-x_0}{x_2-x_0}f(x_2) + \frac{x_2-x}{x_2-x_0}f(x_0).$$

Similarly, since $x_1 < x_0 < x$, we have

$$f(x_0) \le \frac{x_0-x_1}{x-x_1}f(x) + \frac{x-x_0}{x-x_1}f(x_1)$$

which translates to

$$f(x) \ge \frac{x-x_1}{x_0-x_1}f(x_0) - \frac{x-x_0}{x_0-x_1}f(x_1).$$

We have $f(x)$ bounded above and below to the right of x_0 by functions that converge to $f(x_0)$ as x tends to x_0 from the right. Therefore, $f(x)$ is continuous from the right at x_0. By symmetry, $f(x)$ must also be continuous from the left at x_0, and so f is continuous at x_0. Since x_0 is arbitrary, this completes the proof.

(b) Let $f(x)$ be defined on the closed interval $[0,1]$, as $f(x)=0$ for $0<x<1$, $f(0)=1$, and $f(1)=1$. It is easy to see that f is convex on $[0,1]$, but not continuous at $x=0$ or $x=1$..

2.8.3. If $L(\mathbf{a}) \ge \epsilon|\mathbf{a}|+c$ for some $\epsilon>0$, then $L(\mathbf{a}) \to \infty$ as $|\mathbf{a}| \to \infty$, whether or not L is convex.

Conversely, assume without loss of generality that $\mathbf{0} \in \mathcal{A}$, and let $b = L(\mathbf{0})$. Since $L(\mathbf{a}) \to \infty$ as $|\mathbf{a}| \to \infty$, we may find $z>0$ such that $L(\mathbf{a}) \ge b+1$ for all $|\mathbf{a}| \ge z$. Then, from the convexity of L, we have $L(\mathbf{a}) \ge |\mathbf{a}|/z + b$ for $|\mathbf{a}| \ge z$. For $|\mathbf{a}| \le z$, $L(\mathbf{a})$ is bounded below by some number, say c: $L(\mathbf{a}) \ge c$ for $|\mathbf{a}| \le z$. Then, $L(\mathbf{a}) \ge |\mathbf{a}|/z + (c-1)$ for all $\mathbf{a} \in \mathcal{A}$, completing the proof.

2.8.4. Suppose $EZ=\infty$. We are to show that there is no number $a>0$ such that $L(\theta,a) \le EL(\theta,Z)$ for all $\theta>0$. We will show that $e^{-\theta a} > Ee^{-\theta Z}$, or equivalently that $E(1-e^{-\theta(Z-a)}) > 0$, for all θ sufficiently close to zero. First note, since the slope of $1-e^{-x}$ is 1 at the origin, that given any $\alpha<1$ there is a sufficiently small number $B(\alpha)$ such that if $0 \le z \le B(\alpha)$, then $1-e^{-z} \ge \alpha z$. ($B(\alpha)$ is the positive root of the equation $1-e^{-z}=\alpha z$.) Hence,

$$\begin{aligned} E(1-e^{-\theta(Z-a)}) &> E\{(1-e^{-\theta(Z-a)})I(\theta(Z-a) \le B(\alpha))\} \\ &\ge \alpha\theta E\{(Z-a)I((Z-a) \le B(\alpha)/\theta)\}. \end{aligned}$$

Since $EZ=\infty$, this last expectation tends to infinity as θ tends to 0, showing that for θ sufficiently small, the right side of this inequality is positive, and completing the proof.

2.8.5. First we show a special case. Suppose $g(x)$ is a function with nonnegative second derivative, defined on an interval containing 0 and 1 in the interior. Then the Fundamental Theorem of Calculus gives $g(x) = g(0) + \int_0^x g'(y)\,dy$. Applying the same theorem once again to $g'(y)$ gives

$$g(x) = g(0) + \int_0^x [g'(0) + \int_0^y g''(z)\,dz]\,dy = g(0) + xg'(0) + \int_0^x \int_0^y g''(z)\,dz\,dy.$$

If we put $x=1$ in this equation, solve for $g'(0)$, and substitute back into this equation, we find

$$\begin{aligned} g(x) &= g(0) + x[g(1)-g(0)] - x\int_0^1\int_0^y g''(z)\,dz\,dy + \int_0^x\int_0^y g''(z)\,dz\,dy \\ &= xg(1) + (1-x)g(0) - x\int_0^1 g''(z)\,dz + \int_0^x g''(z)(x-z)\,dz \\ &= xg(1) + (1-x)g(0) - x\int_x^1 g''(z)(1-z)\,dz - (1-x)\int_0^x g''(z)z\,dz \\ &\le xg(1) + (1-x)g(0). \end{aligned}$$

(a) and (b): Let $\mathbf{x}$ and $\mathbf{y}$ be elements of S. For $0 \le \alpha \le 1$, let $g(\alpha) = f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})$. Then g satisfies the conditions of the above result with $g'(\alpha) = \dot{f}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})^T(\mathbf{y}-\mathbf{x})$, and, since $\ddot{f}$ is assumed to be nonnegative definite, $g''(\alpha) = (\mathbf{y}-\mathbf{x})^T\ddot{f}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})(\mathbf{y}-\mathbf{x}) \ge 0$. Hence we have $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = g(\alpha) \le \alpha g(1) + (1-\alpha)g(0) = \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$.

For the converse, let $\mathbf{z} \in S$, take $\alpha = 1/2$, $\mathbf{x} = \mathbf{z} + \epsilon\mathbf{a}$ and $\mathbf{y} = \mathbf{z} - \epsilon\mathbf{a}$, where $\mathbf{a}$ is an arbitrary unit vector and $\epsilon > 0$ is sufficiently small so that $\mathbf{x} \in S$ and $\mathbf{y} \in S$. The convexity of f then implies that $f(\mathbf{z}) = f((1/2)\mathbf{x} + (1/2)\mathbf{y}) \le (1/2)[f(\mathbf{x}) + f(\mathbf{y})] = (1/2)[f(\mathbf{z}+\epsilon\mathbf{a}) + f(\mathbf{z}-\epsilon\mathbf{a})]$. This is equivalent to

$$[f(\mathbf{z}+\epsilon\mathbf{a}) - f(\mathbf{z})] - [f(\mathbf{z}-\epsilon\mathbf{a}) - f(\mathbf{z})] \ge 0.$$

Now divide both sides by ϵ^2 and let $\epsilon \to 0$ to find $\mathbf{a}^T\ddot{f}(\mathbf{z})\mathbf{a} \ge 0$. This is true for all unit vectors $\mathbf{a}$ and hence for all vectors $\mathbf{a}$, showing that $\ddot{f}$ is nonnegative definite.

(c) $f(\mathbf{x})$ is convex on S if and only if

(1) $$f(p\mathbf{x} + (1-p)\mathbf{y}) \le pf(\mathbf{x}) + (1-p)f(\mathbf{y}) \quad \text{for all } \mathbf{x} \in S,\ \mathbf{y} \in S \text{ and } 0 \le p \le 1.$$

For $0 \le \alpha \le 1$, let $h(\alpha) = f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})$. The convexity of h for every $\mathbf{x} \in S$ and $\mathbf{y} \in S$ is equivalent to

$$h(p\alpha + (1-p)\beta) \le ph(\alpha) + (1-p)h(\beta)$$

or, equivalently,

(2) $$f[(p\alpha + (1-p)\beta)\mathbf{x} + (1 - p\alpha - (1-p)\beta)\mathbf{y}] \le pf[\alpha\mathbf{x} + (1-\alpha)\mathbf{y}] + (1-p)f[\beta\mathbf{x} + (1-\beta)\mathbf{y}]$$

for all $0 \le p \le 1$. (1) with $\mathbf{x}$ and $\mathbf{y}$ replaced by $\alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ and $\beta\mathbf{x} + (1-\beta)\mathbf{y}$ respectively, gives (2). And (2) with $\alpha = 1$ and $\beta = 0$ gives (1). Thus, (1) and (2) are equivalent.

2.8.6. We compute the matrix of second derivatives of f and show it is nonnegative definite. Then Exercise 2.8.5(b) implies that f is convex.

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{pmatrix} \\ &= p(1-p)\begin{pmatrix} x^{p-2}y^{1-p} & -x^{p-1}y^{-p} \\ -x^{p-1}y^{-p} & x^p y^{-p-1} \end{pmatrix} \\ &= p(1-p)x^{p-2}y^{-p-1}\begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix}.\end{aligned}$$

The coefficient in front of the matrix is positive, so we just check the matrix is nonnegative definite:

$$(a \quad b)\begin{pmatrix} y^2 & -xy \\ -xy & x^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = (ay - bx)^2 \ge 0.$$

2.8.7. Since f is convex, Jensen's inequality implies that $Ef(X,Y) \ge f(EX, EY)$, or equivalently $-EX^pY^{1-p} \ge -(EX)^p(EY)^{1-p}$, which gives Hölder's inequality.

2.8.8. (Note: There is a misprint in the definition of the density $f_X(x)$. The quantity (x^2) should be $\binom{2}{x}$.) The rule δ corresponds to the behavioral rule: if $x = 0$, choose 0 w.p. 1/2 and choose 1/2 w.p. 1/2; if $x = 1$, choose 1/2; if $x = 2$, choose 1/2 w.p. 1/2 and choose 1 w.p. 1/2. Since the loss is convex, we may obtain a better rule by replacing $\delta(X)$ by its expectation: $d(0) = 1/4$, $d(1) = 1/2$, and $d(2) = 3/4$. Since $R(\theta, d_1) = E_\theta(\theta - X/2)^2 = \mathrm{Var}(X|\theta) = \theta(1-\theta)/2$, and $R(\theta, d_2) = E_\theta(\theta - 1/2)^2 = (\theta - 1/2)^2$, we have

$$R(\theta, \delta) = \theta(1-\theta)/4 + (\theta - 1/2)^2/2.$$

Since $d(X) = (X+1)/4$, we have

$$\begin{aligned}R(\theta, d) &= \mathrm{Var}(d(X)|\theta) + \mathrm{Bias}(d(X)|\theta)^2 \\ &= \theta(1-\theta)/8 + (\theta - 1/2)^2/4,\end{aligned}$$

exactly half the value of $R(\theta, \delta)$.

2.8.9. (Note: In the definition of "strictly convex", it should be assumed that $\mathbf{x} \neq \mathbf{y}$.) In the proof of Jensen's inequality, proceed as far as equation (2.13) and consider the case $p_{k+1} > 0$. In this case, we may divide through the inequality by p_{k+1} and define $p'_j = p_j/p_{k+1}$. This gives the inequality

$$f(\mathrm{E}\mathbf{Z}) \leq f(\mathbf{z}) + \sum_1^k p'_j (z_j - \mathrm{E}Z_j) \qquad \text{for all } \mathbf{z} \in S. \tag{3}$$

We clearly have equality at $\mathbf{z} = \mathrm{E}\mathbf{Z} \in S$. We want to show that there is strict inequality at all other points $\mathbf{z} \in S$. Suppose there is equality at some other point $\mathbf{z}' \in S$. Let $\mathbf{z}'' = (1/2)\mathrm{E}\mathbf{Z} + (1/2)\mathbf{z}'$. Then, $\mathbf{z}'' \in S$ and since f is strictly convex, we have

$$\begin{aligned} f(\mathbf{z}'') < (1/2)f(\mathrm{E}\mathbf{Z}) + (1/2)f(\mathbf{z}') &= f(\mathrm{E}\mathbf{Z}) - \sum_1^k p'_j[(1/2)z_j + (1/2)z'_j - \mathrm{E}Z_j] \\ &= f(\mathrm{E}\mathbf{Z}) - \sum_1^k p'_j[z''_j - \mathrm{E}Z_j] \end{aligned}$$

which contradicts (3). Therefore, (3) holds with strict inequality at all points of S other than $\mathrm{E}\mathbf{Z}$. If we now replace $\mathbf{z}$ in (3) with $\mathbf{Z}$ and take expectations, we would have $f(\mathrm{E}\mathbf{Z}) < \mathrm{E}f(\mathbf{Z})$ with strict inequality unless $\mathbf{Z}$ gives its entire mass to the point $\mathrm{E}\mathbf{Z}$.

2.8.10. We prove the more general statement: If $D' \subset D^*$, and D' is essentially complete, and if $S' = \{(R(\theta_1, \delta), \ldots, R(\theta_k, \delta)) : \delta \in D'\}$ is closed, then S is closed from below.

Proof. Let $\mathbf{x} \in \lambda(S)$, i.e. $Q_{\mathbf{x}} \cap \overline{S} = \{\mathbf{x}\}$. We are to show that $\mathbf{x} \in S$. Since $\mathbf{x} \in \overline{S}$, we may find points $\mathbf{x}_1, \mathbf{x}_2, \ldots \in S$, such that $\mathbf{x}_n \to \mathbf{x}$ as $n \to \infty$. Since D' is essentially complete, there are points $\mathbf{x}'_1, \mathbf{x}'_2, \ldots \in S'$, such that $\mathbf{x}'_i \leq \mathbf{x}_i$ for all i. It is sufficient to show that $\mathbf{x}$ is a limit point of the $\mathbf{x}'_n$, since then $\mathbf{x} \in \overline{S'} = S' \subset S$.

Suppose that $\mathbf{x}$ is not a limit point of the $\mathbf{x}_n$. Then there is a $\delta > 0$ such that $|\mathbf{x} - \mathbf{x}'_n| > \delta$ for all n. For n sufficiently large, $|\mathbf{x} - \mathbf{x}_n| < \delta$. Hence, there exist $\mathbf{y} \in S$ on the line joining $\mathbf{x}_n$ and $\mathbf{x}'_n$, such that $|\mathbf{x} - \mathbf{y}_n| = \delta$. Since S is convex, and $\mathbf{x}_n \in S$ and $\mathbf{x}'_n \in S' \subset S$, we have $\mathbf{y}_n \in S$. Since the $\{\mathbf{y}_n\}$ are bounded, there exists a limit point, call it $\mathbf{y}$. Then $\mathbf{y} \in \overline{S}$, and since $\mathbf{x}'_n \leq \mathbf{x}_n \to \mathbf{x}$, we have $\mathbf{y} \leq \mathbf{x}$. i.e. $\mathbf{y} \in Q_{\mathbf{x}}$. Thus, $\mathbf{y} \in Q_{\mathbf{x}} \cap \overline{S}$, and yet $\mathbf{y} \neq \mathbf{x}$ since $d(\mathbf{x}, \mathbf{y}) = \delta$. Contradiction. ∎