1. $f(x) = \left(\prod_{i=1}^{d} x_i \right)^{1/d}$

$$f(\alpha x + (1-\alpha)y) = \left[\prod_{i=1}^{d} (\alpha x_i + (1-\alpha) y_i) \right]^{1/d}$$

$$\alpha f(x) = \alpha \left(\prod_{i=1}^{d} x_i \right)^{1/d}$$

$$(1-\alpha) f(y) = (1-\alpha) \left(\prod_{i=1}^{d} y_i \right)^{1/d}$$

$$\frac{\alpha f(x) + (1-\alpha) f(y)}{f(\alpha x + (1-\alpha) y)} = \frac{\alpha \left(\prod_{i=1}^{d} x_i \right)^{1/d} + (1-\alpha) \left(\prod_{i=1}^{d} y_i \right)^{1/d}}{\left[\prod_{i=1}^{d} (\alpha x_i + (1-\alpha) y_i) \right]^{1/d}}$$

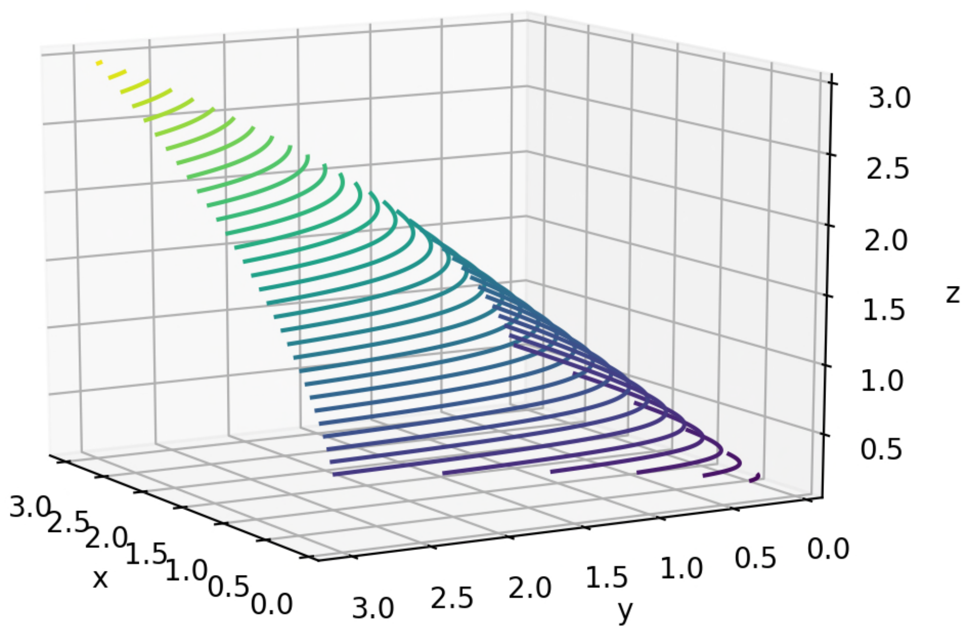
$$= \alpha \left(\prod_{i=1}^{d} \frac{x_i}{\alpha x_i + (1-\alpha) y_i} \right)^{1/d} + (1-\alpha) \left(\prod_{i=1}^{d} \frac{y_i}{\alpha x_i + (1-\alpha) y_i} \right)^{1/d}$$

$$\le \alpha \, \frac{1}{d} \sum_{i=1}^{d} \frac{x_i}{\alpha x_i + (1-\alpha) y_i} + (1-\alpha) \frac{1}{d} \sum_{i=1}^{d} \frac{y_i}{\alpha x_i + 1 - \alpha y_i}$$

$$= \frac{1}{d} \sum_{i=1}^{d} \frac{\alpha x_i + (1-\alpha) y_i}{\alpha x_i + 1 - \alpha y_i}$$

$$= 1$$

$$\Rightarrow \alpha f(x) + (1-\alpha) f(y) \le f(\alpha x + (1-\alpha) y) \Rightarrow \text{concave}$$



3.0
2.5
2.0
1.5
1.0
0.5
z
3.0
2.5
2.0
1.5
1.0
0.5
0.0
x
3.0
2.5
2.0
1.5
1.0
0.5
0.0
y

2. a) $f(x) = \sqrt{x}$

$$f(\alpha x + (1-\alpha)y) = \sqrt{\alpha x + (1-\alpha)y}$$

$$\alpha f(x) = \alpha\sqrt{x} \qquad (1-\alpha)f(y) = (1-\alpha)\sqrt{y}$$

Notice $\quad x + y \geq 2\sqrt{xy}$

$$\Rightarrow \alpha(1-\alpha)x + \alpha(1-\alpha)y \geq 2\alpha(1-\alpha)\sqrt{xy}$$

$$\Rightarrow \alpha x + (1-\alpha)y \geq \alpha^2 x + (1-\alpha)^2 y + 2\alpha(1-\alpha)\sqrt{xy}$$

$$\Rightarrow \left(\sqrt{\alpha x + (1-\alpha)y}\right)^2 \geq \left(\alpha\sqrt{x} + 1-\alpha\sqrt{y}\right)^2$$

$$\Rightarrow f^2(\alpha x + (1-\alpha)y) \geq \left(\alpha f(x) + (1-\alpha)f(y)\right)^2$$

Since $x, y \in (0, \infty)$ both side always greater than 0

$$\Rightarrow f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$$

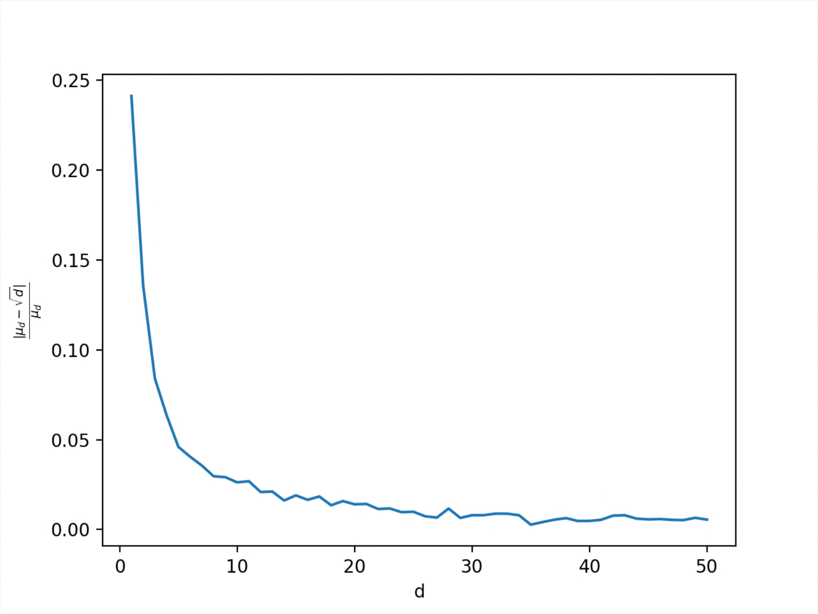
Hence $f(x) = \sqrt{x}$ is concave on $(0, \infty)$

b) $$E\|g\|_2^2 = E[g^T g] = E\sum_{i=1}^{d} g_i^2 = \sum_{i=1}^{d} E g_i^2 = \sum_{i=1}^{d} \left(E g_i^2 - (E g_i)^2\right) = \sum_{i=1}^{d} \sigma^2 = d\sigma^2$$

c) $$E\|g\|_2 = E\sqrt{\|g\|_2^2} \underset{①}{\le} \sqrt{E\|g\|_2^2} = \sigma\sqrt{d}$$

① hold because $f(x)=\sqrt{x}$ is concave, then $E f(x) \le f(EX)$

d)



Observation:

when d increases

$\sigma\sqrt{d}$ is a more precise upper bound for $E\|g\|_2$

3. Notice $\nabla f(z) = \frac{\partial}{\partial z} \|x - z\|_2^2 = 2(z - x)$

and $\nabla f(z_*) = 2(z_* - x)$

$\forall z_0 \in B_2(r)$,
$$\langle \nabla f(z_*), z_0 - z_* \rangle$$
$$= 2 \langle z_* - x, z_0 - z_* \rangle$$
$$= 2\left(\langle z_*, z_0 \rangle - \langle x, z_0 \rangle + \langle z_*, x \rangle - \langle z_*, z_* \rangle \right)$$
$$= 2\left(\frac{r}{\|x\|_2} x^T z_0 - x^T z_0 + \frac{r}{\|x\|_2} x^T x - \frac{r^2}{\|x\|_2^2} x^T x \right)$$
$$= 2\left(\left(\frac{r}{\|x\|_2} - 1\right) x^T z_0 + r\|x\|_2 - r^2 \right)$$
$$= 2\left(\left(\frac{r}{\|x\|_2} - 1\right) x^T z_0 + r(\|x\|_2 - r) \right)$$
$$= 2\left(r(\|x\|_2 - r) - \left(1 - \frac{r}{\|x\|_2}\right) \langle x, z_0 \rangle \right)$$
$$\geq 2\left(r(\|x\|_2 - r) - \left(1 - \frac{r}{\|x\|_2}\right) \|x\|_2 \|z_0\|_2 \right)$$
$$= 2(\|r - z_0\|)(\|x\|_2 - r) \geq 0$$

Since the optimality condition is satisfied.

$z_* = \frac{r}{\|x\|} x$ is indeed the minimizer of

$f(x) = \|x - z\|_2^2$ for all $\|z\|_2 \leq r$.

4. $f(x) = \frac{1}{2}\|x-b\|_2^2 + \lambda\|x\|_1$

$$\nabla f(x) = \frac{\partial}{\partial x}\frac{1}{2}(x-b)^T(x-b) + \frac{\partial}{\partial x}\lambda\|x\|_1$$

$$= \frac{1}{2}\cdot 2(x^T - b^T) + \lambda\frac{\partial}{\partial x}\|x\|_1$$

notice $\frac{\partial}{\partial x_i}\|x\|_1 = \frac{\partial}{\partial x_i}(|x_1| + \dots + |x_d|)$

$$= \begin{cases} 1 & \text{if } x_i > 0 \\ [-1, 1] & \text{if } x_i = 0 \\ -1 & \text{if } x_i < 0 \end{cases}$$

thus $\nabla f(x) = 0 \iff [\nabla f(x)]_i = 0 \quad \forall i = 1, 2, \dots d$

$$[\nabla f(x)]_i = x_i - b_i + \lambda\frac{\partial}{\partial x_i}\|x\|_1 = \begin{cases} x_i - b_i + \lambda & \text{if } x_i > 0 \\ x_i - b_i + [-\lambda, \lambda] & \text{if } x_i = 0 \\ x_i - b_i - \lambda & \text{if } x_i < 0 \end{cases}$$

Using opt condition, Solve for $[\nabla f(x)]_i = 0$ for all i

we get

$$[S_\lambda(b)]_i = x_i = \begin{cases} b_i - \lambda & \text{if } b_i \in [\lambda, \infty) \\ 0 & \text{if } b_i \in [-\lambda, \lambda] \\ b_i + \lambda & \text{if } b_i \in (-\infty, -\lambda] \end{cases}$$