الف) فرم کلی بهینه‌سازی الگوریتم‌های جستجوی خط الگوریتم‌های جستجوی خط (Line Search) به منظور یافتن مقدار بهینه تابع هدف در امتداد یک جهت خاص به کار می‌روند. فرم کلی این بهینه‌سازی به صورت زیر است:

$$x_{k+1} = x_k + \alpha_k d_k$$

که در آن x_k نقطه فعلی و α_k طول پله یا مقدار گام در جهت d_k و d_k جهت جستجو (جهت حرکت) است.
هدف این است که طول گام α_k به گونه‌ای انتخاب شود که مقدار تابع هدف $f(x)$ در امتداد d_k حداقل شود:

$$\alpha_k = \arg\min_{\alpha>0} f(x_k + \alpha d_k)$$

ب) جهت نزول (d_k) به جهتی گفته می‌شود که در آن مقدار تابع هدف $f(x)$ کاهش یابد. این شرط به صورت زیر تعریف می‌شود:

$$\nabla f(x_k)^T d_k < 0$$

پ) برای یافتن طول پله α_k از روش‌های مختلفی استفاده می‌شود، از جمله:
۱. روش دقیق (Exact Line Search): α_k به گونه‌ای محاسبه می‌شود که $f(x_k + \alpha_k d_k)$ دقیقاً کمینه شود. این روش معمولاً مستلزم حل یک مسئله بهینه‌سازی تک‌بعدی است و ممکن است هزینه‌بر باشد.
۲. روش تقریبی (Inexact Line Search): به جای یافتن مقدار دقیق α_k، از شرط‌های تقریبی استفاده می‌شود.
۳. روش ثابت: یک مقدار ثابت برای α_k انتخاب می‌شود. این روش ساده است اما ممکن است ناکارآمد باشد.
۴. روش کاهش تدریجی :(Backtracking) با مقدار اولیه α_0 شروع شده و مقدار α_k به صورت تدریجی کاهش داده می‌شود (مثلاً با ضرب در یک ثابت $\beta < 1$).
۵. روش‌های پیشرفته‌تر استفاده از مشتق دوم: در روش‌های نیوتونی، از اطلاعات مشتق دوم برای انتخاب طول پله استفاده می‌شود:

$$\alpha_k = -\frac{\nabla f(x_k)^T d_k}{d_k^T H_k d_k}$$

که H_k ماتریس هسین تابع است.
۶. روش‌های مبتنی بر یادگیری: در مسائل پیچیده، ممکن است از مدل‌های یادگیری ماشین برای پیش‌بینی طول گام بهینه استفاده شود.

سوال ۱: ب) p یک جهت نزول است اگر و تنها اگر $\nabla f^T(x)p<0$

$$\frac{\partial f}{\partial x_1}=2(x_1+x_2^2)\ ,\ \frac{\partial f}{\partial x_2}=4x_2(x_1+x_2^2)\Rightarrow \nabla f(x)=\begin{pmatrix}\frac{\partial f(x)}{\partial x_1}\\ \frac{\partial f(x)}{\partial x_2}\end{pmatrix}=\begin{pmatrix}2\\4\end{pmatrix}$$

$$\Rightarrow \nabla f^T(x)p=(2\ \ 4)\begin{pmatrix}-1\\-1\end{pmatrix}=-2-4=-6<0 \Rightarrow$$ جهت نزول

$$g(\alpha)=f(x^t+\alpha p^t)\Rightarrow \frac{dg(\alpha^*)}{d\alpha}=\nabla^T f(x^t+\alpha^* p^t)p^t=0\ ,\ x^t+\alpha^* p^t=\begin{pmatrix}x_1^t+\alpha^* p_1^t\\ x_2^t+\alpha^* p_2^t\end{pmatrix}$$

$$\Rightarrow \left(2((x_1^t+\alpha^* p_1^t)+(x_2^t+\alpha^* p_2^t)^2)\quad 4(x_2^t+\alpha^* p_2^t)((x_1^t+\alpha^* p_1^t)+(x_2^t+\alpha^* p_2^t)^2)\right)\begin{pmatrix}p_1^t\\ p_2^t\end{pmatrix}=0$$

$$\Rightarrow \cancel{2}\left(p_1^t+2p_2^t(x_2^t+\alpha^* p_2^t)\right)\left((x_1^t+\alpha^* p_1^t)+(x_2^t+\alpha^* p_2^t)^2\right)=0$$

$$\Rightarrow \begin{cases} p_1^t + 2p_2^t(x_2^t + \alpha^* p_2^t) = 0 \Rightarrow \alpha^* = -\dfrac{p_1^t}{2{p_2^t}^2} - \dfrac{x_2^t}{p_2^t} \\ x_1^t + \alpha^* p_1^t + {x_2^t}^2 + 2\alpha^* x_2^t p_2^t + {\alpha^*}^2 {p_2^t}^2 = 0 \Rightarrow {p_2^t}^2 {\alpha^*}^2 + (p_1^t + 2x_2^t p_2^t)\alpha^* + x_1^t + {x_2^t}^2 = 0 \end{cases}$$

$$\Rightarrow \alpha^* = \frac{-(p_1^t + 2x_2^t p_2^t) \pm \sqrt{\Delta}}{2{p_2^t}^2}, \quad \Delta = (p_1^t + 2x_2^t p_2^t)^2 - 4{p_2^t}^2(x_1^t + {x_2^t}^2)$$

$$p = \begin{pmatrix} -1 \\ -1 \end{pmatrix} : \alpha^* = \frac{1}{2} + x_2^t, \quad \frac{1 + 2x_2^t \pm \sqrt{(1+2x_2^t)^2 - 4(x_1^t + {x_2^t}^2)}}{2}$$

$$x = \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \alpha^* = \frac{3}{2}, \quad \frac{3 \pm \sqrt{5}}{2}$$