Problem 1.

A) $\underset{\sim}{\mathbf{A}}$ is Positive-Definite if $\underset{\sim}{\mathbf{x}}^{\mathrm{T}}\underset{\sim}{\mathbf{A}}\underset{\sim}{\mathbf{x}} > 0$ for non-zero $\underset{\sim}{\mathbf{x}}$.

B)

$$\det(\underset{\sim}{\mathbf{A}}) = \begin{vmatrix} 6 & 2 & -2 \\ 2 & 5 & 0 \\ -2 & 0 & 7 \end{vmatrix}$$

$$= (6-\lambda)(5-\lambda)(7-\lambda) - 2\cdot 2(7-\lambda) - 2(2(5-\lambda))$$

$$= -\lambda^3 + 18\lambda^2 - 99\lambda + 162 = 0$$

$$\Rightarrow (3-\lambda)(6-\lambda)(9-\lambda) = 0$$

$$\Rightarrow \lambda_1 = 9, \lambda_1 = 6, \lambda_1 = 3$$

Since every $\lambda_i > 0$, $\underset{\sim}{\mathbf{A}}$ is Positive-Definite

C)

i) $\det(\underset{\sim}{\mathbf{A}}^1) = |5| = 5 > 0$

ii) $\det(\underset{\sim}{\mathbf{A}}^2) = \begin{vmatrix} 5 & 3 \\ 3 & 5 \end{vmatrix} = 16 > 0$

ii) $\det(\underset{\sim}{\mathbf{A}}^3) = \begin{vmatrix} 5 & 3 & 2 \\ 3 & 5 & 2 \\ 2 & 2 & 6 \end{vmatrix} = 80 > 0$

Since every $\det(\underset{\sim}{\mathbf{A}}^i) > 0$, $\underset{\sim}{\mathbf{A}}$ is Positive-Definite

메모 [kmj1]: 3pts.

메모 [kmj2]: 2pt.

메모 [kmj3]: Using determinant of matrix and accurate calculation: 4pts. -1pt. per error.

메모 [kmj4]: Conclusion: 1pt.

메모 [kmj5]: Using Sylvester's test and accurate calculation: 4pts. -1pt. per error

메모 [kmj6]: Conclusion: 1pt.

Problem 2.

For $f \in C^2$ and f = convex,

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1) \qquad (*)$$

If we use Taylor expansion:

$$f(x_2) \underset{\text{Taylor expansion}}{=} f(x_1) + f'(x_1)(x_2 - x_1) + \frac{1}{2} f''(x_1)(x_2 - x_1)^2 + \cdots$$

$$\underset{\text{using } (*)}{\geq} f(x_1) + f'(x_1)(x_2 - x_1)$$

$$\therefore\ f''(x_1) \geq 0$$

Problem 3.

A)

$$I_n = I_{n+1} + I_{n+2} \qquad (1)$$

$$I_n = \tau I_{n-1} \ \Rightarrow\ I_n = \tau^{n-1} I_1 \qquad (2)$$

Using (1) and (2),

$$I_1 = \tau I_1 + \tau^2 I_1$$

$$\Rightarrow \tau^2 + \tau - 1 = 0 \qquad \therefore \tau = \frac{-1 + \sqrt{5}}{2} \simeq 0.618$$

B) We can increase the search interval at every step.
(or We can increase the interval elimination speed at every step.) .

메모 [kmj7]: Definition of convex function: 4pts.

메모 [kmj8]: Using Taylor expansion and deriving the conclusion: 6pts.

메모 [kmj9]: Any reasonable derivation or the one shown in the lecture not will also be OK.

메모 [kmj10]: Either of the two should be mentioned.

Problem 4.

A)

$$(\mathbf{d}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}} = \left[-\mathbf{g}^{\mathbf{k}} + \beta^{\mathbf{k}-1}\mathbf{d}^{\mathbf{k}-1}\right]^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}$$

$$= -(\mathbf{g}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}} + \beta^{\mathbf{k}-1}\underbrace{\cancel{(\mathbf{d}^{\mathbf{k}-1})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}}}_{\text{by (c)}}$$

$$= -(\mathbf{g}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}$$

B) Recall

$$\underset{\sim}{\mathbf{A}}\mathbf{x}^{k+1} = \underset{\sim}{\mathbf{A}}\left(\mathbf{x}^{k} + \alpha^{k}\mathbf{d}^{\mathbf{k}}\right)$$

$$\Rightarrow \underset{\sim}{\mathbf{A}}\mathbf{d}^{\mathbf{k}} = \frac{1}{\alpha^{k}}\left(\underset{\sim}{\mathbf{A}}\mathbf{x}^{k+1} - \underset{\sim}{\mathbf{A}}\mathbf{x}^{k}\right)$$

$$= \frac{1}{\alpha^{k}}\left(\left(\underset{\sim}{\mathbf{A}}\mathbf{x}^{\mathbf{k}+1} + \mathbf{B}\right) - \left(\underset{\sim}{\mathbf{A}}\mathbf{x}^{\mathbf{k}} + \mathbf{B}\right)\right)$$

$$= \frac{1}{\alpha^{k}}\left(\mathbf{g}^{\mathbf{k}+1} - \mathbf{g}^{\mathbf{k}}\right) \qquad (1)$$

Replacing (1) for $\underset{\sim}{\mathbf{A}}\mathbf{d}^{\mathbf{k}}$ in (a),

$$\beta^{\mathbf{k}} = \frac{(\mathbf{g}^{\mathbf{k}+1})^{\mathbf{T}}\underset{\sim}{\mathbf{A}}(\mathbf{d}^{\mathbf{k}})}{(\mathbf{d}^{\mathbf{k}})^{\mathbf{T}}\underset{\sim}{\mathbf{A}}(\mathbf{d}^{\mathbf{k}})} = \frac{(\mathbf{g}^{\mathbf{k}+1})^{\mathbf{T}}(\mathbf{g}^{\mathbf{k}+1} - \mathbf{g}^{\mathbf{k}})}{(\mathbf{d}^{\mathbf{k}})^{\mathbf{T}}(\mathbf{g}^{\mathbf{k}+1} - \mathbf{g}^{\mathbf{k}})}$$

$$= \frac{(\mathbf{g}^{\mathbf{k}+1})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}+1} - \overbrace{\cancel{(\mathbf{g}^{\mathbf{k}+1})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}}}^{\text{by (e)}}}{\underbrace{\cancel{(\mathbf{d}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}+1}}}_{\text{by (c)}} - \underbrace{(\mathbf{d}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}}_{\text{by (d)}}} = \frac{(\mathbf{g}^{\mathbf{k}+1})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}}{(\mathbf{g}^{\mathbf{k}})^{\mathbf{T}}\mathbf{g}^{\mathbf{k}}}$$

메모 [kmj11]: Everything in this part should be shown in the answer sheet. −1 pt. for each missing term.

5pts.

메모 [kmj12]: Everything in this part should be shown in the answer sheet. −1 pt. for each missing term.

3pts

메모 [kmj13]: Applying (c), (d), and (e): 2pts for each. Replacing (1) for (a): 1pt.

Problem 5.

A)

1^{st} iteration

$$\mathbf{d_0} = -\nabla f(\mathbf{x}_o) = [-2 \quad 2]^{\mathbf{T}}$$

$$f(\alpha) = \mathbf{x}_o + \alpha \mathbf{d_0} = 5\alpha^2 - 8\alpha + 5 = 5\left(\alpha - \frac{4}{5}\right)^2 + \frac{9}{5}$$

$f(\alpha)$ is minimized when $\alpha = 4/5$

$$\therefore \begin{cases} \mathbf{x}^1 = [2.4 \quad 1.6]^{\mathbf{T}} \\ f(\mathbf{x}^1) = 1.8 \end{cases}$$

메모 [kmj14]: 3pts. For the algorithm, 2pts. For the result. Every error will cut 1 pt.

2^{nd} iteration

$$\mathbf{d}^1 = -\nabla f(\mathbf{x}^1) = [-1.2 \quad -1.2]^{\mathbf{T}}$$

$$f(\alpha) = \mathbf{x}^1 + \alpha \mathbf{d}^1 = \frac{9}{5}\alpha^2 - \frac{72}{25}\alpha + \frac{9}{5} = \frac{9}{5}\left(\alpha - \frac{4}{5}\right)^2 + \frac{81}{125}$$

$f(\alpha)$ is minimized when $\alpha = 4/5$

$$\therefore \begin{cases} \mathbf{x}^2 = [1.44 \quad 0.64]^{\mathbf{T}} \\ f(\mathbf{x}^2) = 0.6480 \end{cases}$$

메모 [kmj15]: 3pts. For the algorithm, 2pts. For the result. Every error will cut 1 pt.

B)

1^{st} iteration

$$\mathbf{d_0} = -\nabla f(\mathbf{x}_o) = [-2 \quad 2]^{\mathbf{T}}$$

$$f(\alpha) = \mathbf{x}_o + \alpha \mathbf{d_0} = 5\alpha^2 - 8\alpha + 5 = 5\left(\alpha - \frac{4}{5}\right)^2 + \frac{9}{5}$$

$f(\alpha)$ is minimized when $\alpha = 4/5$

$$\therefore \begin{cases} \mathbf{x}^1 = [2.4 \quad 1.6]^{\mathbf{T}} \\ f(\mathbf{x}^1) = 1.8 \end{cases}$$

메모 [kmj16]: 3pts. For the algorithm, 2pts. For the result. Every error will cut 1 pt.

2^{nd} iteration

$$\mathbf{d}^1 = \mathbf{-g^1} + \beta^0 \mathbf{d}_0 = \begin{bmatrix} -1.2 & -1.2 \end{bmatrix}^{\mathbf{T}} + \beta^{\mathbf{k}} \begin{bmatrix} -2 & 2 \end{bmatrix}^{\mathbf{T}}$$

$$\beta^0 = \frac{\begin{bmatrix} -1.2 & -1.2 \end{bmatrix} \begin{bmatrix} -1.2 \\ -1.2 \end{bmatrix}}{\begin{bmatrix} -2 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix}} = \frac{2.88}{8}$$

$$\mathbf{d}^1 = \begin{bmatrix} -1.92 & -0.48 \end{bmatrix}^{\mathbf{T}}$$

$$f(\alpha) = \frac{144}{125}\alpha^2 - \frac{72}{25}\alpha + \frac{9}{5} = \frac{144}{125}\left(\alpha - \frac{5}{4}\right)^2$$

$f(\alpha)$ is minimized when $\alpha = 5/4$

$$\therefore \left(\begin{array}{l} \mathbf{x}^2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^{\mathbf{T}} \\ f(\mathbf{x}^2) = 0 \end{array} \right.$$

C)

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{1}{2}x_1 & 2x_2 - 2 \end{bmatrix}^{\mathbf{T}} = \begin{bmatrix} 0 & 0 \end{bmatrix}^{\mathbf{T}}$$

$$\rightarrow (x_1^*, x_2^*) = (0,1),\ f^* = 0$$

To check if it is minimum point,

$$\mathbf{H} = \nabla(\nabla f) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix}$$

$$\det(\mathbf{H}^1) = \left| \frac{1}{2} \right| = \frac{1}{2} > 0$$

$$\det(\mathbf{H}^2) = \begin{vmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{vmatrix} = 1 > 0$$

Since every $\det(\mathbf{H}^i) > 0$,

$f^* = 0$ is minimum value (or Global minimum)

메모 [kmj17]: 3pts. For the algorithm, 2pts. For the result. Every error will cut 1 pt.

메모 [kmj18]: Using FONC: 2pts Result: 1pt.

메모 [kmj19]: Checking the P-D of H: 1pt. Concluding f^* to be minimum: 1pt.

D) While SDM searches its minimum value in the steepest descent directions and its convergence is affected by its condition number, CGM uses conjugate directions to guarantee the *n*-iteration convergence. Therefore, the result of SDM didn't reach its minimum due to its convergence condition, while the others did.

메모 [kmj20]: "n-iteration convergence guarantee" and "conjugate direction" should be mentioned. Missing each phrase will deprive you of 3pts., maximum 5pts.

E) The convergence condition of the SDM can be improved by substituting the variables.

Since the condition number is defined as the ratio of Hessian's maximum and minimum eigenvalues,

$$|\mathbf{H}-\lambda\mathbf{I}| = \begin{vmatrix} \frac{1}{2}-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

$$\rightarrow \lambda_{\max} = 2, \lambda_{\min} = \frac{1}{2} \rightarrow \text{ condition number } \xi = \frac{2}{1/2} = 4$$

To change the condition number to become $\xi = 1$, the variable can be substituted as:

$$y_1 = \frac{1}{2}x_1, y_2 = x_2$$

$$\rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \rightarrow \mathbf{x^k} = \mathbf{Dy^k}$$

And when we re-define the problem:

$$\min f(y_1, y_2) = y_1^2 - 2y_2 + y_2^2 + 1$$

$$\text{where } \mathbf{x^k} = \mathbf{Dy^k}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

메모 [kmj21]: "Changing condition number →1 " Should be mentioned. 2pts.

메모 [kmj22]: Actual substitution should be written: 1pt.

메모 [kmj23]: Problem should be re-written. 2pt.