Warmup

Let $f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + 2x_2^2 + 3x_1 - 4x_2$

1. Write f in the form

$$f(x) = \frac{1}{2}x^T Q x + b^T x$$

symmetric

2. Determine if f is positive definite
3. Take a deep breath

1. $f(x_1, x_2) = \frac{1}{2}(x_1\ \ x_2)\begin{pmatrix}4 & 2\\ 2 & 4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + (3\ \ -4)\begin{pmatrix}x_1\\ x_2\end{pmatrix}$

2. (a) Principal minor $4 > 0$

$$\det\begin{pmatrix}4 & 2\\ 2 & 4\end{pmatrix} = 16 - 4 = 12 > 0$$

(b) eigenvalues $\det\begin{pmatrix}4-\lambda & 2\\ 2 & 4-\lambda\end{pmatrix} = 0$

$$\iff (4-\lambda)^2 - 4 = 0$$
$$\iff 16 - 4 - 8\lambda + \lambda^2 = 0$$
$$12 - 8\lambda + \lambda^2 = 0$$
$$(\lambda - 2)(\lambda - 6) = 0$$

eigenvalues $\lambda = 2, 6 > 0$

Q is positive definite

3. Left as an exercise for the reader

When Q is symmetric

$\nabla f(x) = Qx + b$

IMPORTANT: Not true when Q is not symmetric

$\nabla(x^T A x)$: for A not symmetric is $(A + A^T)x$

$\nabla^2 f(x) = Q.$

To find a minimizer, by FONC, we need a minimizer x^* to satisfy

$\nabla f(x) = Qx^* + b = 0$

or $x^* = -Q^{-1}b$

Why is this a minimizer?

By SOSC, $\underset{\nabla^2 f(x)}{\overset{\|}{Q}} > 0$

Gradient Descent

- Steepest
- Fixed stepsize

Algorithm, Given f, starting point x_0

For $k=0,1,2,\ldots$

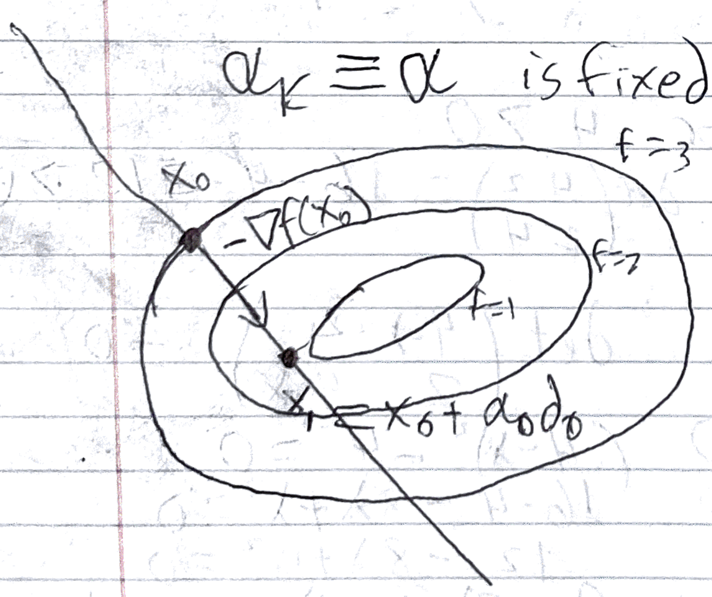
1. Let $d_k = -\nabla f(x_k)$
2. Choose stepsize α_k
3. ~~Let~~ $x_{k+1} = x_k + \alpha_k d_k$

For steepest Descent

$$\alpha_k = \underset{\alpha}{\operatorname{argmin}}\, f(x_k + \alpha\, d_k)$$

For fixed descent

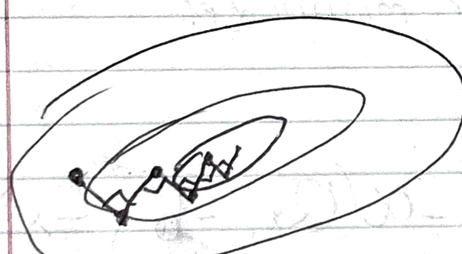
$$\alpha_k \equiv \alpha \text{ is fixed}$$



Theorem:
Consecutive directions of steepest descent are orthogonal.

Proof: Consider points x_k of steepest descent method. We have $x_{k+1} = x_k + \alpha_k d_k$ where α_k minimizes $\phi(\alpha) = f(x_k + \alpha d_k)$

Then by FONC $\phi'(\alpha) = \nabla f(x_k + \alpha d_k)^T d_k = 0$ at $\alpha = \alpha_k$.
Then we see $\nabla f(x_{k+1})^T d_k = 0$
where $d_k = -\nabla f(x_k)$. Therefore consecutive directions are orthogonal.



~~Stepsize Analysis~~

Fact: Steepest Descent is a descent method, i.e. if x_k does not satisfy $\nabla f(x_k) = 0$, then $f(x_{k+1}) < f(x_k)$

Fact: Steepest descent converges to a point where $\nabla f(x_k) = 0$ (under mild conditions, e.g. $f \in C^1$, $f \to \infty$ as $\|x\| \to \infty$)

Fact: This convergence is linear

Fixed Stepsize Analysis for quadratic functions

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

For Quadratic functions of form

$$f(x) = \frac{1}{2} x^T Q x - b^T x$$

$$\nabla f(x) = Qx - b$$

To analyze above method, we consider error $e_k = x_k - x^*$ for minimizer $x^* = Q^{-1} b$

Note $Qx^* = b$

$$e_{k+1} = x_{k+1} - x^* = x_k - \alpha(Qx_k - b) - x^*$$

$$= e_k - \alpha(Qx_k - Qx^*)$$

$$= e_k - \alpha Q e_k$$

$$= (I - \alpha Q) e_k$$

$$= (I - \alpha Q)^{k+1} e_0$$

By spectral Theorem, can write

$e_0 = a_1 v_1 + \dots + a_n v_n$, v_i are eigenvectors of Q with eigenvalue λ_i

$$e_{k+1} = (I - \alpha Q)^{k+1}(a_1 v_1 + \dots + a_n v_n)$$

$$= a_1 (1 - \alpha \lambda_1)^{k+1} v_1 + \dots + (1 - \alpha \lambda_n)^{k+1} v_n$$

For $e_k \to 0$, need all components to go to zero,

~~This occurs when~~

Recall $a^k \to 0$ when $|a|<1$.

Need

$$|1-\alpha\lambda_i|<1 \quad \forall i$$

$$\Leftrightarrow$$

$$-1<1-\alpha\lambda_i<1$$

$$\Leftrightarrow -2<-\alpha\lambda_i<0$$

$$\Leftrightarrow$$

$$\frac{2}{\lambda_i}>\alpha>0 \quad \forall i$$

$$0<\alpha<\frac{2}{\lambda_{max}}$$

Back to warmup problem

Convergence guaranteed when $0<\alpha<\frac{2}{6}=\frac{1}{2}$