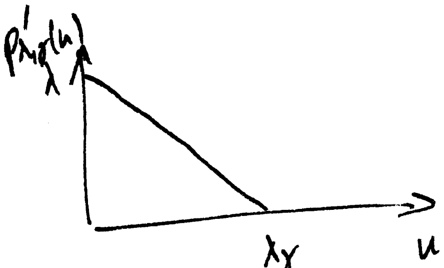
Non-convex penalty

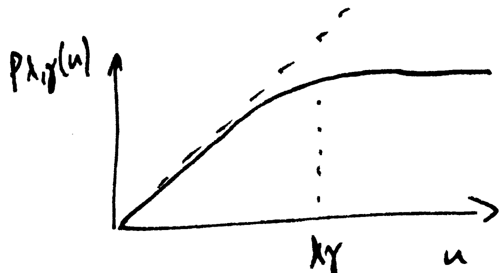
Take a family of non-convex penalty functions $p_{\lambda,\gamma}\ [0,\infty) \to [0,\infty)$ and (attempt to) solve

$$\frac{1}{2n}\|Y - X\beta\|_2^2 + \sum_{k=1}^{p} p_{\lambda,\gamma}(|\beta_k|)$$

A prominent example is the minimax concave penalty (MCP) which takes

$$p'_{\lambda,\gamma}(u) = \left(\lambda - \frac{u}{\gamma}\right)_+ \qquad p_{\lambda,\gamma}(0) = 0$$



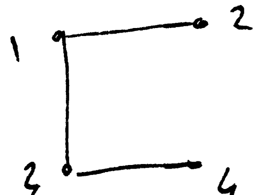


There may be multiple local minima to the penalised objective, which can make optimisation challenging. However coordinate descent often produces reasonable estimators.

Graphs

$V = \{1,2,3,4\}$

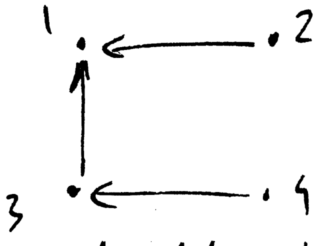
$E = \{(1,2),(2,1)$
$(1,3),(3,1)$
$(3,4),(4,3)\}$



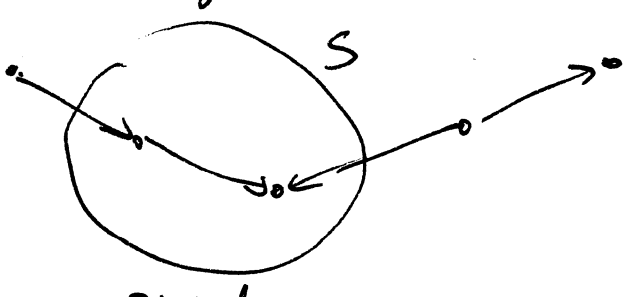
undirected graph

$V = \{1,2,3,4\}$

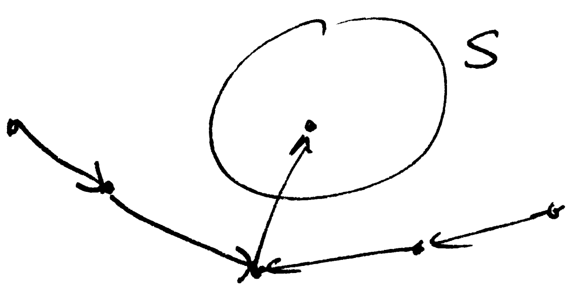
$E = \{(2,1),(3,1),(4,3)\}$



directed graph



Blocked



Not blocked

Prop 28 Given a DAG $\mathcal{G}$ with $V = \{1, \dots, p\}$ say a permutation $\pi : V \to V$ is a <u>topological ordering</u> of the variables if whenever $k \in de(j)$ then $\pi(k) > \pi(j)$.

Every DAG has a topological ordering.

Proof: $p = 1$ ✓ (Induction on p)

Claim: Any DAG has a node with no parents. Pick a node and move to one of its parents (if possible). Then move to one of the new node's parents, and so on. This process must terminate since a DAG has no cycles (no node can be visited more than once). The final node must have no parents; call this the source.

Suppose now $p \geq 2$ and all DAGs of $p-1$ nodes have a topological ordering. Find a source node, w.l.o.g. this is p. Remove p from $\mathcal{G}$ to form $\tilde{\mathcal{G}}$. $\tilde{\mathcal{G}}$ must have a topological ordering $\tilde{\pi}$. But then can take π with $\pi(p) = 1$ and $\pi(k) = \tilde{\pi}(k+1)$ $\forall k \neq p$. □